(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 May 2005 (19.05.2005)

PCT

(10) International Publication Number
WO 2005/045435 A3

(51) International Patent Classification[7]: G01N 33/574, C12Q 1/68

(21) International Application Number:
PCT/EP2004/012462

(22) International Filing Date:
4 November 2004 (04.11.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03025346.2 4 November 2003 (04.11.2003) EP

(71) Applicant *(for DE only)*: **ROCHE DIAGNOSTICS GMBH** [DE/DE]; Sandhofer Strasse 116, 68305 Mannheim (DE).

(71) Applicant *(for all designated States except DE, US)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HAFERLACH, Torsten** [DE/DE]; Springerstrasse 8, 81477 München (DE). **DUGAS, Martin** [DE/DE]; Michael-Fischer-Platz 6, 94469 Deggendorf (DE). **KERN, Wolfgang** [DE/DE]; Hanfelder Strasse 101, 82319 Starnberg (DE). **KOHLMANN, Alexander** [DE/DE]; Schwarzstrasse 14, 92318 Neumarkt (DE). **SCHNITTGER, Susanne** [DE/DE]; Saalburgstrasse 2a, 81375 Munich (DE). **SCHOCH, Claudia** [DE/DE]; Springerstrasse 8, 81477 München (DE).

(74) Common Representative: **ROCHE DIAGNOSTICS GMBH**; c/o Burger Alexander, Patent Department (TR-E), Postfach 11 52, 82372 Penzberg (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
4 August 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR DISTINGUISHING T(11q23)/MLL-POSITIVE LEUKEMIAS FROM T(11q23)MLL NEGATIVE LEUKEMIAS

(57) Abstract: Disclosed is a method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample by determining the expression level of markers, as well as a diagnostic kit and an apparatus containing the markers.

WO 2005/045435 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/012462

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01N33/574 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, WPI Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE BIOSIS [Online] BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 16 November 2002 (2002-11-16), KOHLMANN ALEXANDER ET AL: "Gene Expression Profiles of t(11q23)/MLL Positive ALL and AML." XP002269818 Database accession no. PREV200300356357 abstract & BLOOD, vol. 100, no. 11, 16 November 2002 (2002-11-16), page Abstract No. 308, 44th Annual Meeting of the American Society of Hematology;Philadelphia, PA, USA; December 06-10, 2002 ISSN: 0006-4971 --- -/--	1-27

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 March 2005	09. 06. 2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Thumb, W

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/012462

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TSUTSUMI SHUICHI ET AL: "Two distinct gene expression signatures in pediatric acute lymphoblastic leukemia with MLL rearrangements" CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 63, no. 16, 15 August 2003 (2003-08-15), pages 4882-4887, XP002272402 ISSN: 0008-5472 the whole document	1-27
Y	DATABASE BIOSIS [Online] BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 16 November 2002 (2002-11-16), MOSQUERA-CARO MONICA ET AL: "Heterogeneity of Gene Expression Profiles in MLL-Associated Infant Leukemia: Identification of Distinct Expression Profiles and Novel Therapeutic Targets for Each MLL Translocation Variant." XP002276638 Database accession no. PREV200300357077 abstract & BLOOD, vol. 100, no. 11, 16 November 2002 (2002-11-16), page Abstract No. 2943, 44th Annual Meeting of the American Society of Hematology;Philadelphia, PA, USA; December 06-10, 2002 ISSN: 0006-4971	1-27
Y	ARMSTRONG SCOTT A ET AL: "MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia" NATURE GENETICS, vol. 30, no. 1, January 2002 (2002-01), pages 41-47, XP002276635 ISSN: 1061-4036 the whole document	1-27
Y	ARMSTRONG SCOTT A ET AL: "MLL-rearranged leukemias: Insights from gene expression profiling." SEMINARS IN HEMATOLOGY, vol. 40, no. 4, October 2003 (2003-10), pages 268-273, XP009029193 ISSN: 0037-1963 (ISSN print) the whole document	1-27

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/012462

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ROZOVSKAIA T ET AL: "Upregulation of Meis1 and HoxA9 in acute lymphocytic leukemias with the t(4:11) abnormality" ONCOGENE, vol. 20, no. 7, 15 February 2001 (2001-02-15), pages 874-878, XP002276636 ISSN: 0950-9232 abstract	1-27
Y	SCHOCH CLAUDIA ET AL: "AML with 11q23/MLL abnormalities as defined by the WHO classification: Incidence, partner chromosomes, FAB subtype, age distribution, and prognostic impact in an unselected series of 1897 cytogenetically analyzed AML cases." BLOOD, vol. 102, no. 7, 1 October 2003 (2003-10-01), pages 2395-2402, XP002276637 ISSN: 0006-4971 the whole document	1-27
Y	WO 03/039443 A (DEUTSCHES KREBSFORSCH ;HAFERLACH TORSTEN (DE); EILS ROLAND (DE); K) 15 May 2003 (2003-05-15) the whole document	1-27
Y	KOHLMANN A ET AL: "MOLECULAR CHARACTERIZATION OF ACUTE LEUKEMIAS BY USE OF MICROARRAY TECHNOLOGY" GENES, CHROMOSOMES & CANCER, XX, XX, vol. 37, no. 4, August 2003 (2003-08), pages 396-405, XP008025253 the whole document in particular table 2	1-27
Y	HAFERLACH T ET AL: "The Diagnosis of 14 Specific Subtypes of Leukemia Is Possible Based on Gene Expression Profiles: A Study on 263 Patients with AML, ALL, CML, or CLL" BLOOD, W.B.SAUNDERS COMPAGNY, ORLANDO, FL, US, vol. 100, no. 11, 16 November 2002 (2002-11-16), page 139A, XP002263227 ISSN: 0006-4971 the whole document	1-27
Y	EP 1 043 676 A (WHITEHEAD BIOMEDICAL INST) 11 October 2000 (2000-10-11) the whole document	1-27

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/012462

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GOLUB T R ET AL: "Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 286, no. 5439, 15 October 1999 (1999-10-15), pages 531-537, XP002207658 ISSN: 0036-8075 cited in the application the whole document	1-27
Y	DUGAS M ET AL: "A comprehensive leukemia database: integration of cytogenetics, molecular genetics and microarray data with clinical information, cytomorphology and immunophenotyping" LEUKEMIA, MACMILLAN PRESS LTD, US, vol. 15, no. 12, December 2001 (2001-12), pages 1805-1810, XP002263731 ISSN: 0887-6924 the whole document	1-27
Y	DUGAS MARTIN ET AL: "Impact of integrating clinical and genetic information." IN SILICO BIOLOGY, vol. 2, no. 3, 2002, pages 383-391, XP001179418 ISSN: 1386-6338 (ISSN print) the whole document	1-27
A	ALIZADEH A ET AL: "THE LYMPHOCHIP: A SPECIALIZED CDNA MICROARRAY FOR THE GENOMIC-SCALE ANALYSIS OF GENE EXPRESSION IN NORMAL AND MALIGNANT LYMPHOCYTES" COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY, BIOLOGICAL LABORATORY, COLD SPRING HARBOR, NY, US, vol. 64, no. 1, 1999, pages 71-78, XP001099007 ISSN: 0091-7451 the whole document	1-27
A	DATABASE BIOSIS [Online] BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 16 November 2002 (2002-11-16), KOHLMANN ALEXANDER ET AL: "A Simplified and Partially Automated Target Preparation Method for Gene Expression Profiling." XP002269495 Database accession no. PREV200300367771 abstract	1-27

-/--

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2004/012462

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Article 52 (2)(d) EPC - Presentation of information

The claims were only searched with regards to the underlying method of generating a reference data base for distinguishing leukemia subtypes.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-27 (partially)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-27 (partially)

A method for distinguishing (11q23)/MLL positive leukemias from t(11q23)/MLL negative leukemias, in particular de-novo AML from therapy-related AML, the method comprising determining the expression level of the marker eukaryotic translation elongation factor 1 epsilon 1 (EEF1E1). Use of said marker for the manufacture of a diagnostic. A diagnostic kit containing said marker and an apparatus comprising a reference data bank, wherein the reference data bank is obtainable by determining the expression level of EEF1E1.

---

2. claims: 1-27 (all partially)

Inventions 2-550

Methods for distinguishing t(11q23)/MLL positive leukemias from t(11q23)/MLL negative leukemias, including the subtypes ALL or AML with t(11q23)/MLL, de-novo AML, therapy-related AML, ALL and AML with MLL/t(11;19), ALL with t(4;11), AML with MLL/t(9;11) and AML with MLL/t(6;11) and methods for distinguishing specific subtypes against all other subtypes and against each other, the method comprising determining individually the expression level of the markers listed in table 1, positions 2-50, and in tables 2-7. Use of said markers for the manufacture of diagnostics. Diagnostic kits containing said markers and apparatus comprising a reference data bank, wherein the reference data bank is obtainable by determining the expression levels of said markers.

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	& BLOOD, vol. 100, no. 11, 16 November 2002 (2002-11-16), page Abstract No. 4287, 44th Annual Meeting of the American Society of Hematology;Philadelphia, PA, USA; December 06-10, 2002 ISSN: 0006-4971 -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP2004/012462

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03039443 A	15-05-2003	EP 1308522 A1	07-05-2003
		WO 03039443 A2	15-05-2003
		EP 1470247 A2	27-10-2004
EP 1043676 A	11-10-2000	CA 2304876 A1	09-10-2000
		EP 1043676 A2	11-10-2000
		JP 2001017171 A	23-01-2001
		US 2003017481 A1	23-01-2003
		US 6647341 B1	11-11-2003
		US 2003073083 A1	17-04-2003

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 May 2005 (19.05.2005)

PCT

(10) International Publication Number
WO 2005/045435 A2

(51) International Patent Classification[7]: G01N 33/574, C12Q 1/68

(21) International Application Number: PCT/EP2004/012462

(22) International Filing Date: 4 November 2004 (04.11.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03025346.2 4 November 2003 (04.11.2003) EP

(71) Applicant *(for DE only)*: **ROCHE DIAGNOSTICS GMBH** [DE/DE]; Sandhofer Strasse 116, 68305 Mannheim (DE).

(71) Applicant *(for all designated States except DE, US)*: **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, CH-4070 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HAFERLACH, Torsten** [DE/DE]; Springerstrasse 8, 81477 München (DE). **DUGAS, Martin** [DE/DE]; Michael-Fischer-Platz 6, 94469 Deggendorf (DE). **KERN, Wolfgang** [DE/DE]; Hanfelder Strasse 101, 82319 Starnberg (DE). **KOHLMANN, Alexander** [DE/DE]; Schwarzstrasse 14, 92318 Neumarkt (DE). **SCHNITTGER, Susanne** [DE/DE]; Saalburgstrasse 2a, 81375 Munich (DE). **SCHOCH, Claudia** [DE/DE]; Springerstrasse 8, 81477 München (DE).

(74) Common Representative: **ROCHE DIAGNOSTICS GMBH**; c/o Burger Alexander, Patent Department (TR-E), Postfach 11 52, 82372 Penzberg (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR DISTINGUISHING T(11q23)/MLL-POSITIVE LEUKEMIAS FROM T(11q23)MLL NEGATIVE LEUKEMIAS

(57) Abstract: Disclosed is a method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample by determining the expression level of markers, as well as a diagnostic kit and an apparatus containing the markers.

WO 2005/045435 A2

Method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias

The present invention is directed to a method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias by determining the expression level of selected marker genes.

Leukemias are classified into four different groups or types: acute myeloid (AML), acute lymphatic (ALL), chronic myeloid (CML) and chronic lymphatic leukemia (CLL). Within these groups, several subcategories can be identified further using a panel of standard techniques as described below. These different subcategories in leukemias are associated with varying clinical outcome and therefore are the basis for different treatment strategies. The importance of highly specific classification may be illustrated in detail further for the AML as a very heterogeneous group of diseases. Effort is aimed at identifying biological entities and to distinguish and classify subgroups of AML which are associated with a favorable, intermediate or unfavorable prognosis, respectively. In 1976, the FAB classification was proposed by the French-American-British co-operative group which was based on cytomorphology and cytochemistry in order to separate AML subgroups according to the morphological appearance of blasts in the blood and bone marrow. In addition, it was recognized that genetic abnormalities occurring in the leukemic blast had a major impact on the morphological picture and even more on the prognosis. So far, the karyotype of the leukemic blasts is the most important independent prognostic factor regarding response to therapy as well as survival.

Usually, a combination of methods is necessary to obtain the most important information in leukemia diagnostics: Analysis of the morphology and cytochemistry of bone marrow blasts and peripheral blood cells is necessary to establish the diagnosis. In some cases the addition of immunophenotyping is mandatory to separate very undifferentiated AML from acute lymphoblastic leukemia and CLL. Leukemia subtypes investigated can be diagnosed by cytomorphology alone, only if an expert reviews the smears. However, a genetic analysis based on chromosome analysis, fluorescence in situ hybridization or RT-PCR and immunophenotyping is required in order to assign all cases into the right category. The aim of these techniques besides diagnosis is mainly to determine the

prognosis of the leukemia. A major disadvantage of these methods, however, is that viable cells are necessary as the cells for genetic analysis have to divide in vitro in order to obtain metaphases for the analysis. Another problem is the long time of 72 hours from receipt of the material in the laboratory to obtain the result. Furthermore, great experience in preparation of chromosomes and even more in analyzing the karyotypes is required to obtain the correct result in at least 90% of cases. Using these techniques in combination, hematological malignancies in a first approach are separated into chronic myeloid leukemia (CML), chronic lymphatic (CLL), acute lymphoblastic (ALL), and acute myeloid leukemia (AML). Within the latter three disease entities several prognostically relevant subtypes have been established. As a second approach this further sub-classification is based mainly on genetic abnormalities of the leukemic blasts and clearly is associated with different prognoses.

The sub-classification of leukemias becomes increasingly important to guide therapy. The development of new, specific drugs and treatment approaches requires the identification of specific subtypes that may benefit from a distinct therapeutic protocol and, thus, can improve outcome of distinct subsets of leukemia. For example, the new therapeutic drug (STI571, Imatinib) inhibits the CML specific chimeric tyrosine kinase BCR-ABL generated from the genetic defect observed in CML, the BCR-ABL-rearrangement due to the translocation between chromosomes 9 and 22 (t(9;22) (q34; q11)). In patients treated with this new drug, the therapy response is dramatically higher as compared to all other drugs that had been used so far. Another example is the subtype of acute myeloid leukemia AML M3 and its variant M3v both with karyotype t(15;17)(q22; q11-12). The introduction of a new drug (all-trans retinoic acid - ATRA) has improved the outcome in this subgroup of patient from about 50% to 85 % long-term survivors. As it is mandatory for these patients suffering from these specific leukemia subtypes to be identified as fast as possible so that the best therapy can be applied, diagnostics today must accomplish sub-classification with maximal precision. Not only for these subtypes but also for several other leukemia subtypes different treatment approaches could improve outcome. Therefore, rapid and precise identification of distinct leukemia subtypes is the future goal for diagnostics.

Thus, the technical problem underlying the present invention was to provide means for leukemia diagnostics which overcome at least some of the disadvantages of the prior art diagnostic methods, in particular encompassing the time-consuming and unreliable combination of different methods and which provides a rapid assay to unambiguously distinguish one AML subtype from another, e.g. by genetic analysis.

According to Golub et al. (Science, 1999, 286, 531-7), gene expression profiles can be used for class prediction and discriminating AML from ALL samples. However, for the analysis of acute leukemias the selection of the two different subgroups was performed using exclusively morphologic-phenotypical criteria. This was only descriptive and does not provide deeper insights into the pathogenesis or the underlying biology of the leukemia. The approach reproduces only very basic knowledge of cytomorphology and intends to differentiate classes. The data is not sufficient to predict prognostically relevant cytogenetic aberrations.

Furthermore, the international application WO-A 03/039443 discloses marker genes the expression levels of which are characteristic for certain leukemia, e.g. AML subtypes and additionally discloses methods for differentiating between the subtype of AML cells by determining the expression profile of the disclosed marker genes. However, WO-A 03/039443 does not provide guidance which set of distinct genes discriminate between two subtypes and, as such, can be routineously taken in order to distinguish one AML and/or ALL subtype from another.

The problem is solved by the present invention, which provides a method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample, the method comprising determining the expression level of markers selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7 ,

wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 4, 7, 8, 9, 11, 12, 13, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, 33, 34, 35, 37, 38, 40, 41, 42, 44, 45, 46, 47, 48, 49, and/or 50 of Table 1, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 5, 6, 10, 14, 15, 18, 28, 31, 36, 39, and/or 43 of Table 1,

is indicative for the presence of denovo_AML when denovo_AML is distinguished from therapy-related AML,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 6, 7, 10, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 43, 44, 45, 46, 47, 48, 49, and/or 50 of Table 2, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 5, 8, 9, 11, 12, 14, 24, 28, 33, 41, and/or 42, of Table 2

is indicative for the presence of ALL with t(11q23) when ALL with t(11q23) is distinguished from AML with t(11q23),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22, 25, 26, 27, 28, 29, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 43, 45, 46, 48, 49, and/or 50 of Table 3 and/or

a higher expression of at least one polynucleotide defined by any of the numbers 3, 6, 15, 19, 23, 24, 30, 31, 39, 44, and/or 47, of Table 3

is indicative for the presence of ALL with MLL/t(11;19) when ALL with MLL/t(11;19) is distinguished from AML with MLL/t(11;19)

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 7, 8, 9, 10, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 33, 36, 37, 38, 40, 41, 42, 44, 45, 47, and/or 50 of Table 4, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 5, 6, 11, 12, 14, 19, 26, 32, 34, 35, 39, 43, 46, 48, and/or 49 of Table 4,

is indicative for the presence of ALL with MLL/t(11;19) when ALL with MLL/t(11;19) is distinguished from ALL with MLL/t(4;11),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 19, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, and/or 50 of Table 5, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 10, 15, 16, 20, 22, 32, 33, and/or 42 of Table 5

is indicative for the presence of ALL with MLL/t(9;11) when ALL with MLL/t(9;11) is distinguished from AML with t(11q23),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 4, 8, 13, 14, 16, 21, 22, 23, 24, 29, 30, 31, 36, 37, 38, 39, 44, 48, and/or 49, of Table 6.1, and or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 5, 6, 7, 9, 10, 11, 12, 15, 17, 18, 19, 20, 25, 26, 27, 28, 32, 33, 34, 35, 40, 41, 42, 43, 45, 46, 47, and/or 50 of Table 6.1,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 5, 6, 7, 9, 10, 12, 13, 14, 15, 17, 19, 20, 22, 23, 24, 25, 26, 27, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 48, 49, and/or 50 of Table 6.2, and/or

a higher expression a polynucleotide defined by any of the numbers 1, 2, 3, 4, 8, 11, 16, 18, 21, 28, 30, 31, 35, 36, and/or 47, of Table 6.2

is indicative for the presence of AML with MLL/t(9;11) when AML with MLL/t(9;11) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 5, 6, 9, 10, 17, 18, 19, 21, 22, 23, 26, 27, 28, 31, 32, 34, 36, 37, 39, 41, 42, 44, 46, 47, and/or 49, of Table 6.3, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 4, 7, 8, 11, 12, 13, 14, 15, 16, 20, 24, 25, 29, 30, 33, 35, 38, 40, 43, 45, 48, and/or 50 of Table 6.3

is indicative for the presence of AML with MLL/t(11;19) when AML with MLL/t(11;19) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 2, 7, 8, 16, 17, 18, 22, 33, 34, 35, 36, 48, 49, and/or 50 of Table 7.1, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, and/or 47, of Table 7.1,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from AML with MLL/t(9;11),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 5, 6, 8, 10, 12, 14, 16, 19, 20, 23, 27, 33, 36, 39, 41, 45, 47, 48, 49, of Table 7.2, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 7, 9, 11, 13, 15, 17, 18, 21, 22, 24, 25, 26, 28, 29, 30, 31, 32, 34, 35, 37, 38, 40, 42, 43, 44, 46, and/or 50 of Table 7.2,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from AML with MLL/t(11;19),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 16, 17, 19, 21, 22, 23, 24, 25, 26, 28, 30, 31, 32, 36, 39, 44, 45, 48, 49, of Table 7.3, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 7, 8, 15, 18, 20, 27, 29, 33, 34, 35, 37, 38, 40, 41, 42, 43, 46, 47, and/or 50 of Table 7.3

is indicative for the presence of AML with MLL/t(9;11) when AML with MLL/t(9;11) is distinguished from AML with MLL/t(11;19).

As used herein, the following definitions apply to the above abbreviations:

therapy-related AML (t-AML)

de novo AML: newly existing AML

AML with MLL/t(11;19): AML with (11,19) Translocation

AML with MLL/t(11q23): AML with (11q23) Translocation

AML with MLL/t(6;11): AML with (6;11) Translocation

AML with MLL/t(4;11): AML with (4;11) Translocation

AML with MLL/t(9;11): AML with (9;11) Translocation

As used herein, "all other subtypes" refer to the subtypes of the present invention, i.e. if one subtype is distinguished from "all other subtypes", it is distiguished from all other subtypes contained in the present invention.

According to the present invention, a "sample" means any biological material containing genetic information in the form of nucleic acids or proteins obtainable or obtained from an individual. The sample includes e.g. tissue samples, cell samples, bone marrow and/or body fluids such as blood, saliva, semen. Preferably, the sample is blood or bone marrow, more preferably the sample is bone marrow. The person skilled in the art is aware of methods, how to isolate nucleic acids and proteins from a sample. A general method for isolating and preparing nucleic acids from a sample is outlined in Example 3.

According to the present invention, the term "lower expression" is generally assigned to all by numbers and Affymetrix Id. definable polynucleotides the t-values and fold change (fc) values of which are negative, as indicated in the Tables. Accordingly, the term "higher expression" is generally assigned to all by numbers and Affymetrix Id. definable polynucleotides the t-values and fold change (fc) values of which are positive.

According to the present invention, the term "expression" refers to the process by which mRNA or a polypeptide is produced based on the nucleic acid sequence of a gene, i.e. „expression" also includes the formation of mRNA upon transcription. In accordance with the present invention, the term „determining the expression level" preferably refers to the determination of the level of expression, namely of the markers.

Generally, "marker" refers to any genetically controlled difference which can be used in the genetic analysis of a test versus a control sample, for the purpose of assigning the sample to a defined genotype or phenotype. As used herein, "markers" refer to genes which are differentially expressed in, e.g., different AML subtypes. The markers can be defined by their gene symbol name, their encoded protein name, their transcript identification number (cluster identification number), the data base accession number, public accession number or GenBank identifier or, as done in the present invention, Affymetrix identification number, chromosomal location, UniGene accession number and cluster type, LocusLink accession number (see Examples and Tables).

The Affymetrix identification number (affy id) is accessible for anyone and the person skilled in the art by entering the "gene expression omnibus" internet page of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/geo/). In particular, the affy id's of the polynucleotides used for the method of the present invention are derived from the so-called U133 chip. The sequence data of each identification number can be viewed at http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL96

Generally, the expression level of a marker is determined by the determining the expression of its corresponding "polynucleotide" as described hereinafter.

According to the present invention, the term „polynucleotide" refers, generally, to a DNA, in particular cDNA, or RNA, in particular a cRNA, or a portion thereof or a polypeptide or a portion thereof. In the case of RNA (or cDNA), the polynucleotide is formed upon transcription of a nucleotide sequence which is capable of

expression. The polynucleotide fragments refer to fragments preferably of between at least 8, such as 10, 12, 15 or 18 nucleotides and at least 50, such as 60, 80, 100, 200 or 300 nucleotides in length, or a complementary sequence thereto, representing a consecutive stretch of nucleotides of a gene, cDNA or mRNA. In other terms, polynucleotides include also any fragment (or complementary sequence thereto) of a sequence derived from any of the markers defined above as long as these fragments unambiguously identify the marker.

The determination of the expression level may be effected at the transcriptional or translational level, i.e. at the level of mRNA or at the protein level. Protein fragments such as peptides or polypeptides advantageously comprise between at least 6 and at least 25, such as 30, 40, 80, 100 or 200 consecutive amino acids representative of the corresponding full length protein. Six amino acids are generally recognized as the lowest peptidic stretch giving rise to a linear epitope recognized by an antibody, fragment or derivative thereof. Alternatively, the proteins or fragments thereof may be analysed using nucleic acid molecules specifically binding to three-dimensional structures (aptamers).

Depending on the nature of the polynucleotide or polypeptide, the determination of the expression levels may be effected by a variety of methods. For determining and detecting the expression level, it is preferred in the present invention that the polynucleotide, in particular the cRNA, is labelled.

The labelling of the polynucleotide or a polypeptide can occur by a variety of methods known to the skilled artisan. The label can be fluorescent, chemiluminescent, bioluminescent, radioactive (such as ^{3}H or ^{32}P). The labelling compound can be any labelling compound being suitable for the labelling of polynucleotides and/or polypeptides. Examples include fluorescent dyes, such as fluorescein, dichlorofluorescein, hexachlorofluorescein, BODIPY variants, ROX, tetramethylrhodamin, rhodamin X, Cyanine-2, Cyanine-3, Cyanine-5, Cyanine-7, IRD40, FluorX, Oregon Green, Alexa variants (available e.g. from Molecular Probes or Amersham Biosciences) and the like, biotin or biotinylated nucleotides, digoxigenin, radioisotopes, antibodies, enzymes and receptors. Depending on the type of labelling, the detection is done via fluorescence measurements, conjugation to streptavidin and/or avidin, antigen-antibody- and/or antibody-antibody-interactions, radioactivity measurements, as well as catalytic and/or receptor/ligand

interactions. Suitable methods include the direct labelling (incorporation) method, the amino-modified (amino-allyl) nucleotide method (available e.g. from Ambion), and the primer tagging method (DNA dendrimer labelling, as kit available e.g. from Genisphere). Particularly preferred for the present invention is the use of biotin or biotinylated nucleotides for labelling, with the latter being directly incorporated into, e.g. the cRNA polynucleotide by in vitro transcription.

If the polynucleotide is mRNA, cDNA may be prepared into which a detectable label, as exemplified above, is incorporated. Said detectably labelled cDNA, in single-stranded form, may then be hybridised, preferably under stringent or highly stringent conditions to a panel of single-stranded oligonucleotides representing different genes and affixed to a solid support such as a chip. Upon applying appropriate washing steps, those cDNAs will be detected or quantitatively detected that have a counterpart in the oligonucleotide panel. Various advantageous embodiments of this general method are feasible. For example, the mRNA or the cDNA may be amplified e.g. by polymerase chain reaction, wherein it is preferable, for quantitative assessments, that the number of amplified copies corresponds relative to further amplified mRNAs or cDNAs to the number of mRNAs originally present in the cell. In a preferred embodiment of the present invention, the cDNAs are transcribed into cRNAs prior to the hybridisation step wherein only in the transcription step a label is incorporated into the nucleic acid and wherein the cRNA is employed for hybridisation. Alternatively, the label may be attached subsequent to the transcription step.

Similarly, proteins from a cell or tissue under investigation may be contacted with a panel of aptamers or of antibodies or fragments or derivatives thereof. The antibodies etc. may be affixed to a solid support such as a chip. Binding of proteins indicative of an AML subtype may be verified by binding to a detectably labelled secondary antibody or aptamer. For the labelling of antibodies, it is referred to Harlow and Lane, "Antibodies, a laboratory manual", CSH Press, 1988, Cold Spring Harbor. Specifically, a minimum set of proteins necessary for diagnosis of all AML subtypes may be selected for creation of a protein array system to make diagnosis on a protein lysate of a diagnostic bone marrow sample directly. Protein Array Systems for the detection of specific protein expression profiles already are available (for example: Bio-Plex, BIORAD, München, Germany). For this

application preferably antibodies against the proteins have to be produced and immobilized on a platform e.g. glasslides or microtiterplates. The immobilized antibodies can be labelled with a reactant specific for the certain target proteins as discussed above. The reactants can include enzyme substrates, DNA, receptors, antigens or antibodies to create for example a capture sandwich immunoassay.

For reliably distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias it is useful that the expression of more than one of the above defined markers is determined. As a criterion for the choice of markers, the statistical significance of markers as expressed in q or p values based on the concept of the false discovery rate is determined. In doing so, a measure of statistical significance called the *q* value is associated with each tested feature. The *q* value is similar to the *p* value, except it is a measure of significance in terms of the false discovery rate rather than the false positive rate (Storey JD and Tibshirani R. Proc.Natl.Acad.Sci., 2003, Vol. 100:9440-5.

In a preferred embodiment of the present invention, markers as defined in Tables 1-7 having a p-value of less than 3E-02, more preferred less than 1.5E-04, most preferred less than 1.5E-05, less than 1.5E-06, are measured.

Of the above defined markers, the expression level of at least two, preferably of at least ten, more preferably of at least 25, most preferably of 50 of at least one of the Tables of the markers is determined.

In another preferred embodiment, the expression level of at least 2, of at least 5, of at least 10 out of the markers having the numbers 1 – 10, 1-20, 1-40, 1-50 of at least one of the Tables are measured.

The level of the expression of the „marker", i.e. the expression of the polynucleotide is indicative of the AML subtype of a cell or an organism. The level of expression of a marker or group of markers is measured and is compared with the level of expression of the same marker or the same group of markers from other cells or samples. The comparison may be effected in an actual experiment or in silico. When the expression level also referred to as expression pattern or expression signature (expression profile) is measurably different, there is according to the invention a meaningful difference in the level of expression. Preferably the

difference at least is 5 %, 10% or 20%, more preferred at least 50% or may even be as high as 75% or 100%. More preferred the difference in the level of expression is at least 200%, i.e. two fold, at least 500%, i.e. five fold, or at least 1000%, i.e. 10 fold.

Accordingly, the expression level of markers expressed lower in a first subtype than in at least one second subtype, which differs from the first subtype, is at least 5 %, 10% or 20%, more preferred at least 50% or may even be 75% or 100%, i.e. 2-fold higher, preferably at least 10-fold, more preferably at least 50-fold, and most preferably at least 100-fold lower in the first subtype. On the other hand, the expression level of markers expressed higher in a first subtype than in at least one second subtype, which differs from the first subtype, is at least 5 %, 10% or 20%, more preferred at least 50% or may even be 75% or 100%, i.e. 2-fold higher, preferably at least 10-fold, more preferably at least 50-fold, and most preferably at least 100-fold higher in the first subtype.

In another embodiment of the present invention, the sample is derived from an individual having leukaemia, preferably AML or ALL.

For the method of the present invention it is preferred if the polynucleotide the expression level of which is determined is in form of a transcribed polynucleotide. A particularly preferred transcribed polynucleotide is an mRNA, a cDNA and/or a cRNA, with the latter being preferred. Transcribed polynucleotides are isolated from a sample, reverse transcribed and/or amplified, and labelled, by employing methods well-known the person skilled in the art (see Example 3). In a preferred embodiment of the methods according to the invention, the step of determining the expression profile further comprises amplifying the transcribed polynucleotide.

In order to determine the expression level of the transcribed polynucleotide by the method of the present invention, it is preferred that the method comprises hybridizing the transcribed polynucleotide to a complementary polynucleotide, or a portion thereof, under stringent hybridization conditions, as described hereinafter.

The term "hybridizing" means hybridization under conventional hybridization conditions, preferably under stringent conditions as described, for example, in Sambrook, J., et al., in "Molecular Cloning: A Laboratory Manual" (1989), Eds. J.

Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY and the further definitions provided above. Such conditions are, for example, hybridization in 6x SSC, pH 7.0 / 0.1% SDS at about 45°C for 18-23 hours, followed by a washing step with 2x SSC/0.1% SDS at 50°C. In order to select the stringency, the salt concentration in the washing step can for example be chosen between 2x SSC/0.1% SDS at room temperature for low stringency and 0.2x SSC/0.1% SDS at 50°C for high stringency. In addition, the temperature of the washing step can be varied between room temperature, ca. 22°C, for low stringency, and 65°C to 70° C for high stringency. Also contemplated are polynucleotides that hybridize at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation, preferably of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH2PO4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 mg/ml salmon sperm blocking DNA, followed by washes at 50°C with 1 X SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5x SSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

"Complementary" and "complementarity", respectively, can be described by the percentage, i.e. proportion, of nucleotides which can form base pairs between two polynucleotide strands or within a specific region or domain of the two strands. Generally, complementary nucleotides are, according to the base pairing rules, adenine and thymine (or adenine and uracil), and cytosine and guanine. Complementarity may be partial, in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be a complete or total complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has effects on the efficiency and strength of hybridization between nucleic acid strands.

Two nucleic acid strands are considered to be 100% complementary to each other over a defined length if in a defined region all adenines of a first strand can pair with a thymine (or an uracil) of a second strand, all guanines of a first strand can pair with a cytosine of a second strand, all thymine (or uracils) of a first strand can pair with an adenine of a second strand, and all cytosines of a first strand can pair with a guanine of a second strand, and vice versa. According to the present invention, the degree of complementarity is determined over a stretch of 20, preferably 25, nucleotides, i.e. a 60% complementarity means that within a region of 20 nucleotides of two nucleic acid strands 12 nucleotides of the first strand can base pair with 12 nucleotides of the second strand according to the above ruling, either as a stretch of 12 contiguous nucleotides or interspersed by non-pairing nucleotides, when the two strands are attached to each other over said region of 20 nucleotides. The degree of complementarity can range from at least about 50% to full, i.e. 100% complementarity. Two single nucleic acid strands are said to be "substantially complementary" when they are at least about 80% complementary, preferably about 90% or higher. For carrying out the method of the present invention substantial complementarity is preferred.

Preferred methods for detection and quantification of the amount of polynucleotides, i.e. for the methods according to the invention allowing the determination of the level of expression of a marker, are those described by Sambrook et al. (1989) or real time methods known in the art as the TaqMan® method disclosed in WO92/02638 and the corresponding U.S. 5,210,015, U.S. 5,804,375, U.S. 5,487,972. This method exploits the exonuclease activity of a polymerase to generate a signal. In detail, the (at least one) target nucleic acid component is detected by a process comprising contacting the sample with an oligonucleotide containing a sequence complementary to a region of the target nucleic acid component and a labeled oligonucleotide containing a sequence complementary to a second region of the same target nucleic acid component sequence strand, but not including the nucleic acid sequence defined by the first oligonucleotide, to create a mixture of duplexes during hybridization conditions, wherein the duplexes comprise the target nucleic acid annealed to the first oligonucleotide and to the labeled oligonucleotide such that the 3'-end of the first oligonucleotide is adjacent to the 5'-end of the labeled oligonucleotide. Then this mixture is treated with a template-dependent nucleic acid polymerase having a 5' to 3' nuclease activity under conditions sufficient to permit the 5' to 3' nuclease activity of the polymerase to cleave the annealed, labeled oligonucleotide and

release labeled fragments. The signal generated by the hydrolysis of the labeled oligonucleotide is detected and/ or measured. TaqMan® technology eliminates the need for a solid phase bound reaction complex to be formed and made detectable. Other methods include e.g. fluorescence resonance energy transfer between two adjacently hybridized probes as used in the LightCycler® format described in U.S. 6,174,670.

A preferred protocol if the marker, i.e. the polynucleotide, is in form of a transcribed nucleotide, is described in Example 3, where total RNA is isolated, cDNA and, subsequently, cRNA is synthesized and biotin is incorporated during the transcription reaction. The purified cRNA is applied to commercially available arrays which can be obtained e.g. from Affymetrix. The hybridized cRNA is detected according to the methods described in Example 3. The arrays are produced by photolithography or other methods known to experts skilled in the art e.g. from U.S. 5,445,934, U.S. 5,744,305, U.S. 5,700,637, U.S. 5,945,334 and EP 0 619 321 or EP 0 373 203, or as decribed hereinafter in greater detail.

In another embodiment of the present invention, the polynucleotide or at least one of the polynucleotides is in form of a polypeptide. In another preferred embodiment, the expression level of the polynucleotides or polypeptides is detected using a compound which specifically binds to the polynucleotide of the polypeptide of the present invention.

As used herein, "specifically binding" means that the compound is capable of discriminating between two or more polynucleotides or polypeptides, i.e. it binds to the desired polynucleotide or polypeptide, but essentially does not bind unspecifically to a different polynucleotide or polypeptide.

The compound can be an antibody, or a fragment thereof, an enzyme, a so-called small molecule compound, a protein-scaffold, preferably an anticalin. In a preferred embodiment, the compound specifically binding to the polynucleotide or polypeptide is an antibody, or a fragment thereof.

As used herein, an "antibody" comprises monoclonal antibodies as first described by Köhler and Milstein in Nature 278 (1975), 495-497 as well as polyclonal antibodies, i.e. antibodies contained in a polyclonal antiserum. Monoclonal

antibodies include those produced by transgenic mice. Fragments of antibodies include $F(ab')_2$, Fab and Fv fragments. Derivatives of antibodies include scFvs, chimeric and humanized antibodies. See, for example Harlow and Lane, loc. cit. For the detection of polypeptides using antibodies or fragments thereof, the person skilled in the art is aware of a variety of methods, all of which are included in the present invention. Examples include immunoprecipitation, Western blotting, Enzyme-linked immuno sorbent assay (ELISA), Enzyme-linked immuno sorbent assay (RIA), dissociation-enhanced lanthanide fluoro immuno assay (DELFIA), scintillation proximity assay (SPA). For detection, it is desirable if the antibody is labelled by one of the labelling compounds and methods described supra.

In another preferred embodiment of the present invention, the method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias is carried out on an array.

In general, an "array" or "microarray" refers to a linear or two- or three dimensional arrangement of preferably discrete nucleic acid or polypeptide probes which comprises an intentionally created collection of nucleic acid or polypeptide probes of any length spotted onto a substrate/solid support. The person skilled in the art knows a collection of nucleic acids or polypeptide spotted onto a substrate/solid support also under the term "array". As known to the person skilled in the art, a microarray usually refers to a miniaturised array arrangement, with the probes being attached to a density of at least about 10, 20, 50, 100 nucleic acid molecules referring to different or the same genes per cm^2. Furthermore, where appropriate an array can be referred to as "gene chip". The array itself can have different formats, e.g. libraries of soluble probes or libraries of probes tethered to resin beads, silica chips, or other solid supports.

The process of array fabrication is well-known to the person skilled in the art. In the following, the process for preparing a nucleic acid array is described. Commonly, the process comprises preparing a glass (or other) slide (e.g. chemical treatment of the glass to enhance binding of the nucleic acid probes to the glass surface), obtaining DNA sequences representing genes of a genome of interest, and spotting sequences these sequences of interest onto glass slide. Sequences of interest can be obtained via creating a cDNA library from an mRNA source or by

using publicly available databases, such as GeneBank, to annotate the sequence information of custom cDNA libraries or to identify cDNA clones from previously prepared libraries. Generally, it is recommendable to amplify obtained sequences by PCR in order to have sufficient amounts of DNA to print on the array. The liquid containing the amplified probes can be deposited on the array by using a set of microspotting pins. Ideally, the amount deposited should be uniform. The process can further include UV-crosslinking in order to enhance immobilization of the probes on the array.

In a preferred embodiment, the array is a high density oligonucleotide (oligo) array using a light-directed chemical synthesis process, employing the so-called photolithography technology. Unlike common cDNA arrays, oligo arrays (according to the Affymetrix technology) use a single-dye technology. Given the sequence information of the markers, the sequence can be synthesized directly onto the array, thus, bypassing the need for physical intermediates, such as PCR products, required for making cDNA arrays. For this purpose, the marker, or partial sequences thereof, can be represented by 14 to 20 features, preferably by less than 14 features, more preferably less than 10 features, even more preferably by 6 features or less, with each feature being a short sequence of nucleotides (oligonucleotide), which is a perfect match (PM) to a segment of the respective gene. The PM oligonucleotide are paired with mismatch (MM) oligonucleotides which have a single mismatch at the central base of the nucleotide and are used as "controls". The chip exposure sites are defined by masks and are deprotected by the use of light, followed by a chemical coupling step resulting in the synthesis of one nucleotide. The masking, light deprotection, and coupling process can then be repeated to synthesize the next nucleotide, until the nucleotide chain is of the specified length.

Advantageously, the method of the present invention is carried out in a robotics system including robotic plating and a robotic liquid transfer system, e.g. using microfluidics, i.e. channelled structured.

A particular preferred method according to the present invention is as follows:

1. Obtaining a sample, e.g. bone marrow or peripheral blood aliquots, from a patient having AML or ALL
2. Extracting RNA, preferably mRNA, from the sample
3. Reverse transcribing the RNA into cDNA

4. In vitro transcribing the cDNA into cRNA
5. Fragmenting the cRNA
6. Hybridizing the fragmented cRNA on standard microarrays
7. Determining hybridization

In another embodiment, the present invention is directed to the use of at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7 for the manufacturing of a diagnostic for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias. The use of the present invention is particularly advantageous for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in an individual having AML or ALL. The use of said markers for diagnosis of t(11q23)/MLL-positive leukemias and t(11q23)/MLL negative leukemias, preferably based on microarray technology, offers the following advantages: (1) more rapid and more precise diagnosis, (2) easy to use in laboratories without specialized experience, (3) abolishes the requirement for analyzing viable cells for chromosome analysis (transport problem), and (4) very experienced hematologists for cytomorphology and cytochemistry, immunophenotyping as well as cytogeneticists and molecularbiologists are no longer required.

Accordingly, the present invention refers to a diagnostic kit containing at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7 for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias, in combination with suitable auxiliaries. Suitable auxiliaries, as used herein, include buffers, enzymes, labelling compounds, and the like. In a preferred embodiment, the marker contained in the kit is a nucleic acid molecule which is capable of hybridizing to the mRNA corresponding to at least one marker of the present invention. Preferably, the at least one nucleic acid molecule is attached to a solid support, e.g. a polystyrene microtiter dish, nitrocellulose membrane, glass surface or to non-immobilized particles in solution.

In another preferred embodiment, the diagnostic kit contains at least one reference for a t(11q23)/MLL-positive leukemia and/or for a t(11q23)/MLL negative leukemia. As used herein, the reference can be a sample or a data bank.

In another embodiment, the present invention is directed to an apparatus for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample, containing a reference data bank obtainable by comprising

(a) compiling a gene expression profile of a patient sample by determining the expression level at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7, and

(b) classifying the gene expression profile by means of a machine learning algorithm.

According to the present invention, the "machine learning algorithm" is a computational-based prediction methodology, also known to the person skilled in the art as "classifier", employed for characterizing a gene expression profile. The signals corresponding to a certain expression level which are obtained by the microarray hybridization are subjected to the algorithm in order to classify the expression profile. Supervised learning involves "training" a classifier to recognize the distinctions among classes and then "testing" the accuracy of the classifier on an independent test set. For new, unknown sample the classifier shall predict into which class the sample belongs.

Preferably, the machine learning algorithm is selected from the group consisting of Weighted Voting, K-Nearest Neighbors, Decision Tree Induction, Support Vector Machines (SVM), and Feed-Forward Neural Networks. Most preferably, the machine learning algorithm is Support Vector Machine, such as polynomial kernel and Gaussian Radial Basis Function-kernel SVM models.

The classification accuracy of a given gene list for a set of microarray experiments is preferably estimated using Support Vector Machines (SVM), because there is evidence that SVM-based prediction slightly outperforms other classification techniques like k-Nearest Neighbors (k-NN). The LIBSVM software package version 2.36 was used (SVM-type: C-SVC, linear kernel (http://www.csie.ntu.edu.tw/~cjlin/libsvm/)). The skilled artisan is furthermore referred to Brown et al., Proc.Natl.Acad.Sci., 2000; 97: 262-267, Furey et al., Bioinformatics. 2000; 16: 906-914, and Vapnik V. Statistical Learning Theory. New York: Wiley, 1998.

In detail, the classification accuracy of a given gene list for a set of microarray experiments can be estimated using Support Vector Machines (SVM) as supervised learning technique. Generally, SVMs are trained using differentially expressed genes which were identified on a subset of the data and then this trained model is employed to assign new samples to those trained groups from a second and different data set. Differentially expressed genes were identified applying ANOVA and t-test-statistics (Welch t-test). Based on identified distinct gene expression signatures respective training sets consisting of 2/3 of cases and test sets with 1/3 of cases to assess classification accuracies are designated. Assignment of cases to training and test set is randomized and balanced by diagnosis. Based on the training set a Support Vector Machine (SVM) model is built.

According to the present invention, the apparent accuracy, i.e. the overall rate of correct predictions of the complete data set was estimated by 10fold cross validation. This means that the data set was divided into 10 approximately equally sized subsets, an SVM-model was trained for 9 subsets and predictions were generated for the remaining subset. This training and prediction process was repeated 10 times to include predictions for each subset. Subsequently the data set was split into a training set, consisting of two thirds of the samples, and a test set with the remaining one third. Apparent accuracy for the training set was estimated by 10fold cross validation (analogous to apparent accuracy for complete set). A SVM-model of the training set was built to predict diagnosis in the independent test set, thereby estimating true accuracy of the prediction model. This prediction approach was applied both for overall classification (multi-class) and binary classification (diagnosis X => yes or no). For the latter, sensitivity and specificity were calculated:

Sensitivity = (number of positive samples predicted)/(number of true positives)

Specificity = (number of negative samples predicted)/(number of true negatives)

In a preferred embodiment, the reference data bank is backed up on a computational data memory chip which can be inserted in as well as removed from the apparatus of the present invention, e.g. like an interchangeable module, in order to use another data memory chip containing a different reference data bank.

The apparatus of the present invention containing a desired reference data bank can be used in a way such that an unknown sample is, first, subjected to gene expression profiling, e.g. by microarray analysis in a manner as described supra or in the art, and the expression level data obtained by the analysis are, second, fed into the apparatus and compared with the data of the reference data bank obtainable by the above method. For this purpose, the apparatus suitably contains a device for entering the expression level of the data, for example a control panel such as a keyboard. The results, whether and how the data of the unknown sample fit into the reference data bank can be made visible on a provided monitor or display screen and, if desired, printed out on an incorporated of connected printer.

Alternatively, the apparatus of the present invention is equipped with particular appliances suitable for detecting and measuring the expression profile data and, subsequently, proceeding with the comparison with the reference data bank. In this embodiment, the apparatus of the present invention can contain a gripper arm and/or a tray which takes up the microarray containing the hybridized nucleic acids.

In another embodiment, the present invention refers to a reference data bank for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample obtainable by comprising

(a) compiling a gene expression profile of a patient sample by determining the expression level of at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7, and

(b) classifying the gene expression profile by means of a machine learning algorithm.

Preferably, the reference data bank is backed up and/or contained in a computational memory data chip.

The invention is further illustrated in the following table and examples, without limiting the scope of the invention:

TABLES 1-7

Tables 1-7 show AML subtype analysis of t(11q23)/MLL-positive leukemias and t(11q23)/MLL negative leukemias. The analysed markers are ordered according to their q- and p values, beginning with the lowest q- and p values.
For convenience and a better understanding, Tables 1 to 7 are accompanied with explanatory tables (Table 1A to 7A) where the numbering and the Affymetrix Id are further defined by other parameters, e.g. gene bank accession number.

EXAMPLES

Example 1: General experimental design of the invention and results

Rearrangements of the MLL gene occur in acute lymphoblastic and acute myeloid leukemias (ALL, AML). Recent microarray studies report that t(11q23)/MLL positive leukemias demonstrate specific gene expression patterns. However, less is known both about the impact of various MLL partner genes and the transcriptome of de novo versus therapy-related MLL leukemias. Out of a series of n=195 acute leukemias, analyzed by U133 set microarrays (Affymetrix), we addressed the following questions: (i) identification of MLL versus non-MLL rearranged gene patterns, (ii) discrimination of MLL positive AML versus ALL, (iii) analysis of t(9;11) versus other partner genes in AML, and (iv) identification of gene signatures of therapy-related cases (t-AML) compared to de novo AML. When compared to various subtypes of acute leukemias, t(11q23)/MLL positive cases can be predicted with high accuracies. Support vector machine (SVM) based subtype stratification accurately identifies all 48 MLL cases compared to ALL with t(9;22) (n=23), t(8;14) (n=13), precursor T-ALL (n=23), or AML with t(8;21) (n=25), t(15;17) (n=20), inv(16) (n=25), inv(3) (n=18). This is mainly due to a common overexpression of HOXA family members (HOXA7, HOXA9, HOXA10) and TALE family genes (PBX3, MEIS1) in MLL cases. Secondly, a large number of genes separates MLL positive samples according to the lineage they are derived from. B-lineage commitment in ALL with t(11q23)/MLL (n=17) can be illustrated by expression of PAX5 and downstream genes (CD19, IGHM, BLNK, CD79A) repressing the transcription of non-lymphoid genes and by simultaneously

activating the expression of B-lineage-specific genes. Moreover, this finding can be confirmed when restricted to a stringent comparison of t(11;19) positive ALL versus t(11;19) positive AML cases. We next aimed at identifying signatures correlated with different MLL partner genes. Within t(11q23)/AML t(9;11) positive cases (n=19) were compared to non-t(9;11) positive samples (n=12), and also more detailed to t(6;11) (n=3) and t(11;19) cases (n=4). Neither supervised nor unsupervised analyses of our data revealed that expression signatures are influenced by the different translocation partners. This is an unexpected result but however correlates with the observation of no differences in clinical outcome with respect to varying partner genes (Schoch et al., Blood 2003, in press). Finally, our cohort of t(11q23)/MLL AML samples comprised both de novo AML (n=21) and t-AML (n=12). A specific pattern of genes suggests that there are distinct signatures correlated with t-AML cases. Differing transcriptomes may explain in part the even more unfavorable outcome of this AML subgroup. Genes with higher expression in therapy-related compared with de novo cases were involved in DNA repair, cell proliferation, and cell cycle regulation. Taken together, distinct gene expression profiles can be observed in t(11q23)/MLL positive acute leukemias. Both cell lineage background and t-AML characteristics but not partner genes contribute to fundamental changes in gene expression despite a common underlying genetic aberration.

Example 2: General materials, methods and definitions of functional annotations

The methods section contains both information on statistical analyses used for identification of differentially expressed genes and detailed annotation data of identified microarray probesets.

Affymetrix Probeset Annotation

All annotation data of GeneChip® arrays are extracted from the NetAffx™ Analysis Center (internet website: www.affymetrix.com). Files for U133 set arrays, including U133A and U133B microarrays are derived from the June 2003 release. The original publication refers to: Liu G, Loraine AE, Shigeta R, Cline M, Cheng J, Valmeekam V, Sun S, Kulp D, Siani-Rose MA. NetAffx: Affymetrix probesets and annotations. Nucleic Acids Res. 2003;31(1):82-6.

The sequence data are omitted due to their large size, and because they do not change, whereas the annotation data are updated periodically, for example new information on chromomal location and functional annotation of the respective gene products. Sequence data are available for download in the NetAffx Download Center (www.affymetrix.com)

Data fields:
In the following section, the content of each field of the data files are described. Microarray probesets, for example found to be differentially expressed between different types of leukemia samples are further described by additional information. The fields are of the following types:

1. GeneChip Array Information
2. Probe Design Information
3. Public Domain and Genomic References

1. GeneChip Array Information

HG-U133 ProbeSet_ID:
HG-U133 ProbeSet_ID describes the probe set identifier. Examples are: 200007_at, 200011_s_at, 200012_x_at.

GeneChip:
The description of the GeneChip probe array name where the respective probeset is represented. Examples are: Affymetrix Human Genome U133A Array or Affymetrix Human Genome U133B Array.

2. Probe Design Information

Sequence Type:
The Sequence Type indicates whether the sequence is an Exemplar, Consensus or Control sequence. An Exemplar is a single nucleotide sequence taken directly from a public database. This sequence could be an mRNA or EST. A Consensus sequence, is a nucleotide sequence assembled by Affymetrix, based on one or more sequence taken from a public database.

Transcript ID:
The cluster identification number with a sub-cluster identifier appended.

Sequence Derived From:
The accession number of the single sequence, or representative sequence on which the probe set is based. Refer to the "Sequence Source" field to determine the database used.

Sequence ID:
For Exemplar sequences: Public accession number or GenBank identifier. For Consensus sequences: Affymetrix identification number or public accession number.

Sequence Source:
The database from which the sequence used to design this probe set was taken. Examples are: GenBank®, RefSeq, UniGene, TIGR (annotations from The Institute for Genomic Research).

3. Public Domain and Genomic References

Most of the data in this section come from LocusLink and UniGene databases, and are annotations of the reference sequence on which the probe set is modeled.

Gene Symbol and Title:
A gene symbol and a short title, when one is available. Such symbols are assigned by different organizations for different species. Affymetrix annotational data come from the UniGene record. There is no indication which species-specific databank was used, but some of the possibilities include for example HUGO: The Human Genome Organization.

MapLocation:
The map location describes the chromosomal location when one is available.

Unigene_Accession:
UniGene accession number and cluster type. Cluster type can be "full length" or "est", or "---" if unknown.

LocusLink:
This information represents the LocusLink accession number.

Full Length Ref. Sequences:
Indicates the references to multiple sequences in RefSeq. The field contains the ID and description for each entry, and there can be multiple entries per probeSet.

Example 3: Sample preparation, processing and data analysis

Method 1:

Microarray analyses were performed utilizing the GeneChip® System (Affymetrix, Santa Clara, USA). Hybridization target preparations were performed according to recommended protocols (Affymetrix Technical Manual). In detail, at time of diagnosis, mononuclear cells were purified by Ficoll-Hypaque density centrifugation. They had been lysed immediately in RLT buffer (Qiagen, Hilden, Germany), frozen, and stored at -80°C from 1 week to 38 months. For gene expression profiling cell lysates of the leukemia samples were thawed, homogenized (QIAshredder, Qiagen), and total RNA was extracted (RNeasy Mini Kit, Qiagen). Subsequently, 5-10 µg total RNA isolated from 1 x 10^7 cells was used as starting material for cDNA synthesis with oligo[$(dT)_{24}$T7promotor]$_{65}$ primer (cDNA Synthesis System, Roche Applied Science, Mannheim, Germany). cDNA products were purified by phenol/chlorophorm/IAA extraction (Ambion, Austin, USA) and acetate/ethanol-precipitated overnight. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides were incorporated during the following *in vitro* transcription reaction (Enzo BioArray HighYield RNA Transcript Labeling Kit, Enzo Diagnostics). After quantification by spectrophotometric measurements and 260/280 absorbance values assessment for quality control of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 µg cRNA was fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2/500 mM potassium acetate/150 mM magnesium acetate) and added to the hybridization cocktail sufficient for five hybridizations on standard GeneChip microarrays (300 µl final volume). Washing and staining of the probe arrays was performed according to the recommended Fluidics Station protocol (EukGE-WS2v4). Affymetrix Microarray Suite software (version 5.0.1) extracted fluorescence signal intensities from each

feature on the microarrays as detected by confocal laser scanning according to the manufacturer's recommendations.

Expression analysis quality assessment parameters included visiual array inspection of the scanned image for the presence of image artifacts and correct grid alignment for the identification of distinct probe cells as well as both low 3'/5' ratio of housekeeping controls (mean: 1.90 for GAPDH) and high percentage of detection calls (mean: 46.3% present called genes). The 3' to 5' ratio of GAPDH probesets can be used to assess RNA sample and assay quality. Signal values of the 3' probe sets for GAPDH are compared to the Signal values of the corresponding 5' probe set. The ratio of the 3' probe set to the 5' probe set is generally no more than 3.0. A high 3' to 5' ratio may indicate degraded RNA or inefficient synthesis of ds cDNA or biotinylated cRNA (GeneChip® Expression Analysis Technical Manual, www.affymetrix.com). Detection calls are used to determine whether the transcript of a gene is detected (present) or undetected (absent) and were calculated using default parameters of the Microarray Analysis Suite MAS 5.0 software package.

Method 2:

Bone marrow (BM) aspirates are taken at the time of the initial diagnostic biopsy and remaining material is immediately lysed in RLT buffer (Qiagen), frozen and stored at -80 C until preparation for gene expression analysis. For microarray analysis the GeneChip System (Affymetrix, Santa Clara, CA, USA) is used. The targets for GeneChip analysis are prepared according to the current Expression Analysis. Briefly, frozen lysates of the leukemia samples are thawed, homogenized (QIAshredder, Qiagen) and total RNA extracted (RNeasy Mini Kit, Qiagen).Normally 10 ug total RNA isolated from 1 x 107 cells is used as starting material in the subsequent cDNA-Synthesis using Oligo-dT-T7-Promotor Primer (cDNA synthesis Kit, Roche Molecular Biochemicals). The cDNA is purified by phenol-chlorophorm extraction and precipitated with 100% Ethanol over night. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides are incorporated during the in vitro transcription reaction (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 ug are fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridization cocktail sufficient for 5

hybridizations on standard GeneChip microarrays. Before expression profiling Test3 Probe Arrays (Affymetrix) are chosen for monitoring of the integrity of the cRNA. Only labeled cRNA-cocktails which showed a ratio of the messured intensity of the 3' to the 5' end of the GAPDH gene less than 3.0 are selected for subsequent hybridization on HG-U133 probe arrays (Affymetrix). Washing and staining the Probe arrays is performed as described (siehe Affymetrix-Original-Literatur (LOCKHART und LIPSHUTZ). The Affymetrix software (Microarray Suite, Version 4.0.1) extracted fluorescence intensities from each element on the arrays as detected by confocal laser scanning according to the manufacturers recommendations.

Table 1

1. One-Versus-All (OVA)

1.1 de novo MLL versus therapy-related MLL

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	213907_at	EEF1E1	-1.66	7.44E-06	1.70E-01	-1.05	-5.78	6p24.3-p25.1
2	234260_at		3.44	1.06E-05	1.70E-01	0.91	5.25	
3	232663_s_at		2.97	3.90E-05	2.05E-01	0.98	5.12	
4	206180_x_at	MGC2474	-1.58	4.43E-05	2.05E-01	-0.91	-5.01	16p11.2
5	235513_at		3.28	2.66E-05	2.05E-01	0.87	4.95	
6	238970_at		2.92	2.80E-05	2.05E-01	0.86	4.92	
7	221053_s_at	TDRKH	-1.80	5.63E-05	2.05E-01	-0.89	-4.92	1q21
8	231534_at	CDC2	-2.93	2.58E-04	2.59E-01	-1.04	-4.91	10q21.1
9	208565_at	MC5R	-2.25	1.80E-04	2.59E-01	-0.91	-4.75	18p11.2
10	239897_at	BTF	1.81	4.69E-05	2.05E-01	0.82	4.73	6q22-q23
11	205080_at	RARB	-1.45	6.16E-05	2.05E-01	-0.84	-4.73	3p24
12	230964_at		-2.35	2.58E-04	2.59E-01	-0.92	-4.67	
13	244245_at		-2.08	6.38E-05	2.05E-01	-0.82	-4.66	
14	223251_s_at	ANKRD10	1.64	1.15E-04	2.59E-01	0.84	4.64	13q33.3
15	207715_at	CRYGB	2.66	1.27E-04	2.59E-01	0.81	4.51	2q33-q35
16	234858_at		-1.76	3.62E-04	2.70E-01	-0.88	-4.50	
17	205362_s_at	PFDN4	-1.52	1.56E-04	2.59E-01	-0.79	-4.44	20q13
18	209065_at	UQCRB	1.43	1.09E-04	2.59E-01	0.77	4.43	8q22
19	204993_at	GNAZ	-1.52	1.71E-04	2.59E-01	-0.79	-4.42	22q11.22
20	204826_at	CCNF	-2.16	5.91E-04	2.92E-01	-0.92	-4.42	16p13.3
21	225345_s_at		-2.36	2.46E-04	2.59E-01	-0.81	-4.40	
22	215115_x_at	NTRK3	-1.72	4.21E-04	2.71E-01	-0.85	-4.38	15q25
23	207596_at	PRO2176	-1.78	1.56E-04	2.59E-01	-0.77	-4.38	5q21.1
24	219558_at	FLJ20986	-1.60	1.93E-04	2.59E-01	-0.77	-4.33	3q29
25	210087_s_at	MPZL1	-1.89	1.90E-04	2.59E-01	-0.76	-4.32	1q23.2
26	218685_s_at	SMUG1	-1.97	3.53E-04	2.70E-01	-0.80	-4.30	12q13.11-q13.3
27	230662_at	LOC149603	-2.26	3.41E-04	2.70E-01	-0.80	-4.30	1q42.13
28	206086_x_at	HFE	-1.63	2.06E-04	2.59E-01	-0.76	-4.29	6p21.3
29	203408_s_at	SATB1	2.24	1.78E-04	2.59E-01	0.75	4.27	3p23
30	225491_at		-2.90	7.60E-04	2.96E-01	-0.88	-4.26	
31	227657_at	KIAA1214	2.63	1.86E-04	2.59E-01	0.75	4.26	4q31.1
32	209437_s_at	SPON1	-2.43	4.47E-04	2.71E-01	-0.80	-4.25	11p15.2
33	221701_s_at	FLJ12541	-1.47	4.58E-04	2.71E-01	-0.80	-4.25	15q22.33
34	211904_x_at	RAD52	-1.77	2.26E-04	2.59E-01	-0.75	-4.25	12p13-p12.2
35	210714_at	R3HDM	-1.68	2.24E-04	2.59E-01	-0.75	-4.25	2q21.2
36	237376_at		2.55	1.87E-04	2.59E-01	0.74	4.24	
37	220398_at	MGC4170	-1.49	6.27E-04	2.96E-01	-0.83	-4.23	12q23.3
38	210913_at	CDH20	-1.66	4.00E-04	2.71E-01	-0.78	-4.23	18q22-q23
39	214093_s_at		1.69	1.94E-04	2.59E-01	0.74	4.23	
40	215201_at		3.98	2.81E-04	2.59E-01	0.76	4.21	
41	229478_x_at	BIVM	4.32	3.44E-04	2.70E-01	0.79	4.20	13q32-q33.1

42	212022_s_at	MKI67	-2.17	6.33E-04	2.96E-01	-0.82	-4.20	10q25-qter
43	239851_at		1.95	2.30E-04	2.59E-01	0.73	4.20	
44	212020_s_at	MKI67	-1.74	3.89E-04	2.71E-01	-0.77	-4.19	10q25-qter
45	241106_at		-1.64	2.69E-04	2.59E-01	-0.74	-4.19	
46	209946_at	VEGFC	-1.66	4.84E-04	2.71E-01	-0.78	-4.17	4q34.1-q34.3
47	203019_x_at	SSX2IP	-1.70	6.75E-04	2.96E-01	-0.81	-4.17	
48	223661_at		-1.63	7.19E-04	2.96E-01	-0.82	-4.17	
49	201026_at	IF2	-1.53	7.16E-04	2.96E-01	-0.81	-4.16	2p11.1-q11.1
50	219917_at	FLJ23024	-1.52	2.91E-04	2.59E-01	-0.73	-4.15	4p15.2

Table 2

One-Versus-All (OVA)

2. ALL with t(11q23) versus AML with t(11q23)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	211404_s_at	APLP2	-5.84	1.06E-17	2.66E-13	-2.13	-14.47	11q24
2	208702_x_at	APLP2	-8.03	5.57E-15	4.65E-11	-2.15	-13.55	11q24
3	214875_x_at	APLP2	-7.60	8.73E-14	2.62E-10	-2.04	-12.55	11q24
4	200742_s_at	CLN2	-4.09	2.09E-15	2.61E-11	-1.85	-12.49	11p15
5	41220_at	MSF	2.92	1.48E-11	9.24E-09	2.01	12.24	17q25
6	217800_s_at	NDFIP1	-11.26	9.43E-14	2.62E-10	-1.92	-12.18	5q31.3
7	201858_s_at	PRG1	-2.99	1.80E-14	8.59E-11	-1.75	-11.86	10q22.1
8	225703_at	KIAA1545	3.83	4.42E-10	1.10E-07	2.08	11.58	12q24.33
9	226496_at	FLJ22611	7.62	2.43E-11	1.25E-08	1.87	11.56	9p12
10	221969_at	PAX5	24.83	3.08E-09	5.04E-07	2.49	11.52	9p13
11	244876_at		4.77	8.26E-10	1.81E-07	2.13	11.51	
12	225775_at		3.98	4.62E-12	4.13E-09	1.79	11.49	
13	204122_at	TYROBP	-8.95	8.67E-13	1.55E-09	-1.84	-11.42	19q13.1
14	212207_at	KIAA1025	4.05	8.85E-10	1.89E-07	2.10	11.41	12q24.22
15	200743_s_at	CLN2	-2.84	1.39E-14	8.59E-11	-1.63	-11.25	11p15
16	223120_at	MGC1314	-3.99	2.13E-13	4.85E-10	-1.69	-11.17	6q24
17	205639_at	AOAH	-21.09	2.97E-12	3.23E-09	-1.82	-11.02	7p14-p12
18	219013_at	GALNT11	-6.79	7.02E-13	1.35E-09	-1.69	-11.00	7q34-q36
19	206111_at	RNASE2	-5.13	2.06E-14	8.59E-11	-1.59	-10.98	14q24-q31
20	227853_at		-5.82	2.49E-14	8.89E-11	-1.57	-10.90	
21	210314_x_at	TNFSF13	-6.45	1.62E-13	4.06E-10	-1.58	-10.76	17p13.1
22	209500_x_at	TNFSF13	-5.50	9.48E-13	1.58E-09	-1.61	-10.63	17p13.1
23	222422_s_at	NDFIP1	-10.17	1.66E-12	2.30E-09	-1.63	-10.61	5q31.3
24	230015_at		8.96	6.79E-09	8.76E-07	2.09	10.55	
25	214181_x_at	LST1	-7.39	4.21E-12	4.11E-09	-1.66	-10.54	6p21.3
26	225563_at	LOC255967	4.38	1.62E-09	3.09E-07	1.85	10.51	13q12.13
27	203799_at	BIMLEC	-4.99	1.28E-12	2.00E-09	-1.57	-10.45	2q24.2
28	217979_at	NET-6	10.39	8.52E-09	1.06E-06	2.06	10.40	7p21.1
29	211581_x_at	LST1	-5.53	2.34E-12	2.79E-09	-1.58	-10.36	6p21.3
30	213116_at	NEK3	-5.36	2.50E-12	2.85E-09	-1.57	-10.33	13q14.13
31	200975_at	PPT1	-3.29	2.33E-13	4.86E-10	-1.50	-10.32	1p32
32	229215_at	ASCL2	-8.81	5.20E-12	4.33E-09	-1.60	-10.32	11p15.5
33	243756_at		5.60	5.34E-09	7.31E-07	1.90	10.27	
34	211474_s_at	SERPINB6	-5.40	4.90E-12	4.22E-09	-1.57	-10.22	6p25
35	211582_x_at	LST1	-6.23	6.12E-12	4.94E-09	-1.57	-10.18	6p21.3
36	214574_x_at	LST1	-6.06	1.12E-11	7.76E-09	-1.57	-10.07	6p21.3
37	218942_at	FLJ22055	-6.28	1.50E-12	2.20E-09	-1.48	-10.05	12q13.13
38	202788_at	MAPKAPK3	-2.83	4.43E-12	4.11E-09	-1.51	-10.00	3p21.3
39	202382_s_at	GNPI	-13.13	2.27E-11	1.25E-08	-1.58	-9.94	5q21
40	215633_x_at	LST1	-6.93	2.49E-11	1.25E-08	-1.59	-9.94	6p21.3
41	201874_at	MPZL1	2.41	9.12E-10	1.92E-07	1.63	9.92	1q23.2

42	203796_s_at	BCL7A	6.53	1.59E-08	1.73E-06	1.94	9.91 12q24.13
43	210629_x_at	LST1	-4.91	1.05E-11	7.57E-09	-1.51	-9.89 6p21.3
44	200661_at	PPGB	-6.17	1.06E-11	7.57E-09	-1.49	-9.79 20q13.1
45	218404_at	SNX10	-5.74	4.30E-12	4.11E-09	-1.44	-9.75 7p15.2
46	200871_s_at	PSAP	-5.72	4.02E-11	1.73E-08	-1.56	-9.75 10q21-q22
47	201494_at	PRCP	-3.52	2.38E-11	1.25E-08	-1.51	-9.73 11q14
48	235033_at		-3.55	9.71E-12	7.37E-09	-1.46	-9.71
49	201201_at	CSTB	-4.03	1.84E-12	2.42E-09	-1.41	-9.69 21q22.3
50	216041_x_at	GRN	-7.38	2.95E-11	1.39E-08	-1.51	-9.66 17q21.32

Table 3

One-Versus-All (OVA)

ALL with MLL/t(11;19) versus AML with MLL/t(11;19)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	201413_at	HSD17B4	-7.61	2.70E-06	7.06E-02	-9.15	-25.03	5q21
2	218361_at	FLJ10687	-9.54	4.62E-05	1.63E-01	-7.19	-18.80	1q21.2
3	225590_at	POSH	5.81	4.41E-06	7.06E-02	5.56	15.71	4q32.3
4	228624_at	FLJ11155	-10.67	1.76E-05	1.15E-01	-5.64	-15.58	4q32.1
5	203253_s_at	KIAA0433	-3.08	2.42E-04	1.82E-01	-6.28	-15.53	5q21.1
6	225703_at	KIAA1545	5.00	2.95E-04	1.89E-01	5.66	14.15	12q24.33
7	212516_at	CENTD2	-3.79	4.57E-04	2.36E-01	-5.75	-13.85	11q13.2
8	208621_s_at	VIL2	6.47	1.74E-04	1.80E-01	5.30	13.79	6q25.2-q26
9	213468_at	ERCC2	-5.21	1.80E-05	1.15E-01	-4.68	-13.14	19q13.3
10	217337_at		-5.26	1.97E-04	1.81E-01	-4.89	-12.83	
11	224918_x_at	MGST1	-50.45	1.00E-03	2.59E-01	-6.06	-12.78	12p12.3-p12.1
12	218383_at	C14orf94	-2.96	3.54E-05	1.59E-01	-4.60	-12.77	14q11.2
13	223120_at	MGC1314	-5.12	3.45E-04	2.13E-01	-4.92	-12.56	6q24
14	203672_x_at	TPMT	-3.00	1.59E-05	1.15E-01	-4.42	-12.52	6p22.3
15	210396_s_at		2.98	1.20E-04	1.74E-01	4.61	12.42	
16	201231_s_at	ENO1	-3.31	5.36E-05	1.63E-01	-4.49	-12.39	1p36.3-p36.2
17	216574_s_at	RPE	-10.32	1.69E-04	1.80E-01	-4.50	-12.02	2q32-q33.3
18	210644_s_at	LAIR1	-2.89	3.80E-05	1.59E-01	-4.29	-12.00	19q13.4
19	200099_s_at - HG-U133A		1.18	3.57E-04	2.13E-01	4.61	11.89	
20	209623_at	MCCC2	-2.11	2.56E-04	1.82E-01	-4.28	-11.34	5q12-q13
21	203517_at	MTX2	-3.97	5.87E-05	1.63E-01	-3.98	-11.12	2q31.2
22	225214_at		-4.76	7.36E-04	2.42E-01	-4.42	-11.01	
23	217234_s_at	VIL2	8.46	6.60E-04	2.38E-01	4.32	10.89	6q25.2-q26
24	212651_at	RHOBTB1	7.61	1.36E-03	2.62E-01	4.76	10.84	10q21.2
25	200971_s_at	SERP1	-1.63	3.97E-05	1.59E-01	-3.79	-10.71	3q25.1
26	212513_s_at	VDU1	-2.32	7.98E-05	1.63E-01	-3.78	-10.56	1p31.1
27	200901_s_at	M6PR	-4.17	1.33E-03	2.61E-01	-4.45	-10.46	12p13
28	208967_s_at	AK2	-3.51	9.68E-05	1.63E-01	-3.73	-10.38	1p34
29	203573_s_at	RABGGTA	-2.30	1.51E-03	2.65E-01	-4.42	-10.25	14q11.2
30	38269_at	PRKD2	5.82	1.16E-03	2.59E-01	4.18	10.17	19q13.2
31	214373_at	PPP4R2	2.59	8.51E-05	1.63E-01	3.62	10.14	3q29
32	213589_s_at	LOC284208	-31.42	1.99E-03	2.80E-01	-4.51	-9.91	17q25.3
33	231736_x_at	MGST1	-37.89	2.13E-03	2.80E-01	-4.65	-9.88	12p12.3-p12.1
34	218073_s_at	FLJ10407	-1.87	8.17E-05	1.63E-01	-3.49	-9.81	1p32.3
35	229645_at		18.23	2.11E-03	2.80E-01	4.51	9.80	
36	227711_at	FLJ32942	-14.47	1.73E-03	2.65E-01	-4.16	-9.72	12q13.13
37	209421_at	MSH2	-2.48	6.05E-04	2.36E-01	-3.67	-9.64	2p22-p21
38	225008_at	MGC34646	-4.27	2.32E-04	1.82E-01	-3.49	-9.58	8q12.1
39	239978_at		2.15	4.19E-04	2.32E-01	3.56	9.55	
40	227296_at	LOC113655	-4.14	7.60E-05	1.63E-01	-3.37	-9.54	8q24.3
41	208702_x_at	APLP2	-14.39	1.67E-03	2.65E-01	-4.00	-9.53	11q24

42	202246_s_at	CDK4	-2.90	8.01E-05	1.63E-01	-3.37	-9.52 12q14
43	215767_at		-7.77	8.70E-05	1.63E-01	-3.34	-9.44
44	231431_s_at		4.09	1.14E-03	2.59E-01	3.69	9.31
45	211033_s_at	PEX7	-2.18	9.22E-05	1.63E-01	-3.26	-9.22 6q21-q22.2
46	225510_at		-6.39	2.27E-03	2.81E-01	-3.94	-9.07
47	208881_x_at	IDI1	3.98	2.88E-04	1.89E-01	3.28	9.03 10p15.3
48	203518_at	CHS1	-4.04	1.04E-04	1.67E-01	-3.19	-9.02 1q42.1-q42.2
49	201121_s_at	PGRMC1	-1.30	1.26E-03	2.61E-01	-3.56	-9.01 Xq22-q24
50	205246_at	PEX13	-2.21	9.27E-04	2.59E-01	-3.46	-8.98 2p14-p16

Table 4

One-Versus-All (OVA)

ALL with MLL/t(11;19) versus ALL with MLL/t(4;11)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	213908_at		5.01	1.36E-04	1.77E-01	2.19	8.15	
2	221355_at	CHRNG	2.67	2.18E-06	6.23E-02	1.80	7.43	2q33-q34
3	228180_at		1.66	2.98E-06	6.23E-02	1.78	7.32	
4	213932_x_at	HLA-A	1.34	7.32E-06	1.02E-01	1.69	6.93	6p21.3
5	209732_at	CLECSF2	2.75	3.16E-04	1.79E-01	1.85	6.90	12p13-p12
6	231904_at	U2AF1	1.85	1.62E-05	1.03E-01	1.66	6.78	21q22.3
7	208837_at	P24B	-1.84	1.41E-05	1.03E-01	-1.61	-6.51	15q24-q25
8	208945_s_at	BECN1	-2.40	2.01E-05	1.05E-01	-1.63	-6.46	17q21
9	201063_at	RCN1	-3.00	1.60E-05	1.03E-01	-1.54	-6.32	11p13
10	205708_s_at		-2.11	1.72E-05	1.03E-01	-1.55	-6.31	
11	239615_at		2.26	1.62E-04	1.77E-01	1.60	6.27	
12	238714_at		2.00	1.51E-04	1.77E-01	1.59	6.27	
13	225563_at	LOC255967	-1.76	3.14E-05	1.41E-01	-1.53	-6.26	13q12.13
14	215339_at	NKTR	2.96	2.08E-03	2.05E-01	1.73	5.99	3p23-p21
15	212076_at	MLL	-2.89	3.77E-05	1.41E-01	-1.48	-5.95	11q23
16	212080_at		-3.38	4.99E-05	1.41E-01	-1.50	-5.91	
17	203573_s_at	RABGGTA	-2.06	1.45E-04	1.77E-01	-1.47	-5.88	14q11.2
18	212516_at	CENTD2	-1.68	3.77E-05	1.41E-01	-1.41	-5.78	11q13.2
19	227444_at		1.99	7.55E-05	1.45E-01	1.42	5.78	
20	222875_at	DDX33	-2.10	5.73E-05	1.41E-01	-1.44	-5.75	17p13.2
21	223461_at	LOC51256	-1.88	4.24E-05	1.41E-01	-1.40	-5.73	6p23
22	223109_at	CLONE24922	-7.42	9.58E-05	1.52E-01	-1.54	-5.71	9q34.13
23	212656_at	TSFM	-1.96	5.51E-05	1.41E-01	-1.38	-5.68	12q13-q14
24	218911_at	GAS41	-2.09	5.40E-05	1.41E-01	-1.38	-5.68	12q13-q15
25	222573_s_at	SAV1	-3.35	6.20E-05	1.43E-01	-1.38	-5.66	14q13-q23
26	209684_at	RIN2	3.42	3.70E-04	1.79E-01	1.45	5.66	
27	215967_s_at	LY9	-7.58	5.39E-05	1.41E-01	-1.36	-5.58	1q21.3-q22
28	210487_at	DNTT	-7.67	7.64E-05	1.45E-01	-1.39	-5.57	10q23-q24
29	214845_s_at	CALU	-1.85	6.50E-05	1.43E-01	-1.33	-5.49	7q32
30	231747_at	CYSLTR1	-2.57	7.03E-05	1.45E-01	-1.33	-5.45	Xq13.2-21.1
31	220083_x_at	UCHL5	-1.72	1.87E-04	1.79E-01	-1.34	-5.42	1q32
32	200715_x_at	RPL13A	1.17	9.54E-05	1.52E-01	1.34	5.42	19q13.3
33	205977_s_at	EPHA1	-2.47	9.36E-05	1.52E-01	-1.31	-5.37	7q32-q36
34	241642_x_at		2.01	7.98E-04	2.01E-01	1.40	5.37	
35	234043_at		3.63	3.44E-03	2.29E-01	1.54	5.33	
36	219933_at	GLRX2	-2.35	9.83E-05	1.52E-01	-1.29	-5.28	1q31.2-q31.3
37	208666_s_at	ST13	-1.81	9.44E-05	1.52E-01	-1.28	-5.27	22q13.2
38	201796_s_at	VARS2	-3.67	1.06E-04	1.59E-01	-1.28	-5.25	6p21.3
39	240873_x_at	DAB2	2.23	5.49E-04	1.99E-01	1.33	5.23	5p13
40	202613_at	CTPS	-3.86	1.52E-04	1.77E-01	-1.29	-5.18	1p34.1
41	203677_s_at	TARBP2	-1.78	1.26E-04	1.76E-01	-1.27	-5.17	12q12-q13

42	203023_at	HSPC111	-2.98	1.26E-04	1.76E-01	-1.25	-5.17	5q35.3
43	203733_at	MYLE	1.76	1.66E-04	1.77E-01	1.30	5.16	16p13.2
44	221085_at	TNFSF15	-2.62	1.31E-04	1.77E-01	-1.24	-5.12	9q32
45	211150_s_at	DLAT	-6.43	1.73E-04	1.77E-01	-1.27	-5.10	11q23.1
46	239448_at		2.53	3.61E-04	1.79E-01	1.26	5.07	
47	217337_at		-4.49	2.24E-04	1.79E-01	-1.29	-5.07	
48	236859_at	RUNX2	6.95	1.17E-02	3.09E-01	1.84	5.06	6p21
49	233105_at		1.50	3.31E-04	1.79E-01	1.25	5.05	
50	222052_at		-1.75	1.53E-04	1.77E-01	-1.22	-5.04	

Table 5

One-Versus-All (OVA)

AML with MLL/t(9;11) versus AML with t(11q23)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	235865_at		-2.10	1.01E-04	8.00E-01	-0.85	-4.65	
2	226676_at	EHZF	2.61	9.68E-05	8.00E-01	0.81	4.52	18q11.1
3	238161_at		-1.69	1.77E-04	8.00E-01	-0.81	-4.41	
4	222260_at	PDPK1	-1.73	5.67E-04	8.00E-01	-0.88	-4.34	16p13.3
5	241258_at		-1.97	3.66E-04	8.00E-01	-0.82	-4.31	
6	219602_s_at	FLJ23403	-1.65	5.65E-04	8.00E-01	-0.80	-4.16	18p11.21
7	238353_at		-1.38	5.61E-04	8.00E-01	-0.76	-4.05	
8	244475_at		-1.50	9.22E-04	8.00E-01	-0.80	-4.02	
9	229388_at	LOC118491	-2.13	9.36E-04	8.00E-01	-0.78	-3.98	10q22.2
10	239268_at		2.38	5.01E-04	8.00E-01	0.73	3.97	
11	234836_at		-2.03	6.67E-04	8.00E-01	-0.72	-3.90	
12	244290_at		-2.30	8.04E-04	8.00E-01	-0.72	-3.88	
13	230438_at	TBX15	-1.48	7.55E-04	8.00E-01	-0.72	-3.87	1p11.1
14	237354_at		-1.70	1.21E-03	8.00E-01	-0.76	-3.86	
15	213693_s_at	MUC1	3.43	6.22E-04	8.00E-01	0.70	3.85	1q21
16	204548_at	STAR	3.14	7.78E-04	8.00E-01	0.72	3.85	8p11.2
17	228431_at	FLJ11236	-2.20	8.41E-04	8.00E-01	-0.71	-3.84	
18	236846_at		-1.62	1.55E-03	8.00E-01	-0.78	-3.83	
19	240653_at		-1.94	6.49E-04	8.00E-01	-0.69	-3.82	
20	222773_s_at	GALNT12	2.67	6.97E-04	8.00E-01	0.68	3.80	9q22.33
21	207526_s_at	IL1RL1	-1.59	7.50E-04	8.00E-01	-0.68	-3.79	2q12
22	202564_x_at	ARL2	1.86	9.57E-04	8.00E-01	0.70	3.76	11q13
23	223603_at	ZNF179	-1.64	1.17E-03	8.00E-01	-0.70	-3.75	17p11.2
24	234625_at		-1.53	1.18E-03	8.00E-01	-0.70	-3.74	
25	230772_at		-1.46	1.16E-03	8.00E-01	-0.70	-3.73	
26	230953_at		-1.59	1.41E-03	8.00E-01	-0.71	-3.71	
27	232575_at	PCA3	-2.10	1.48E-03	8.00E-01	-0.70	-3.69	9q21-q22
28	231388_at		-2.01	1.52E-03	8.00E-01	-0.70	-3.68	
29	208076_at	HIST1H4D	-1.54	1.77E-03	8.00E-01	-0.72	-3.68	6p21.3
30	219181_at	LIPG	-1.98	1.83E-03	8.00E-01	-0.72	-3.67	18q21.1
31	241920_x_at	FLJ21439	-1.75	1.51E-03	8.00E-01	-0.69	-3.66	15q14
32	244194_at		-1.51	1.40E-03	8.00E-01	-0.67	-3.63	
33	226677_at	EHZF	2.47	1.10E-03	8.00E-01	0.65	3.63	18q11.1
34	201050_at	PLD3	5.31	1.86E-03	8.00E-01	0.76	3.63	19q13.13
35	233083_at		-2.47	2.88E-03	8.00E-01	-0.76	-3.61	
36	237314_at	MGC26778	-2.04	1.90E-03	8.00E-01	-0.69	-3.60	10p12.1
37	241976_at	TCEA3	-1.70	1.85E-03	8.00E-01	-0.68	-3.59	1p36.11
38	207333_at	NMBR	-2.61	2.65E-03	8.00E-01	-0.72	-3.57	6q21-qter
39	230175_s_at	ESDN	-1.66	1.91E-03	8.00E-01	-0.67	-3.55	3q12.1
40	237718_at	EIF4E	-2.96	2.52E-03	8.00E-01	-0.70	-3.54	4q21-q25
41	208344_x_at	IFNA13	-2.05	2.50E-03	8.00E-01	-0.69	-3.53	9p22

42	240032_at		-1.82	3.92E-03	8.00E-01	-0.78	-3.51	
43	205429_s_at	MPP6	3.47	1.62E-03	8.00E-01	0.64	3.51	7p15
44	222860_s_at	SCDGF-B	-1.90	3.12E-03	8.00E-01	-0.71	-3.50	11q22.3
45	216764_at		-1.74	2.69E-03	8.00E-01	-0.68	-3.49	
46	241176_at		-2.82	3.94E-03	8.00E-01	-0.75	-3.48	
47	243173_at	LOC55954	-2.03	2.29E-03	8.00E-01	-0.65	-3.47	22cen-q12.3
48	209560_s_at	DLK1	-1.58	2.50E-03	8.00E-01	-0.66	-3.47	14q32
49	200033_at - HG-U133A	DDX5	-1.23	1.77E-03	8.00E-01	-0.62	-3.46	17q21
50	233003_at		-1.48	1.88E-03	8.00E-01	-0.62	-3.45	

Table 6

One-Versus-All (OVA)

6.1 AML with MLL/t(6;11) versus rest

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	213721_at	SOX2	-6.39	3.74E-09	1.57E-04	-1.77	-9.00	3q26.3-q27
2	217506_at		1.92	9.10E-07	2.29E-03	1.83	8.93	
3	207056_s_at	SLC4A8	2.72	7.59E-07	2.29E-03	1.75	8.62	12q13
4	202233_s_at	UQCRH	-1.51	2.13E-08	4.46E-04	-1.64	-8.28	1p33
5	217655_at		4.49	6.19E-03	2.60E-01	2.39	8.04	
6	243109_at	FLJ11175	1.88	1.48E-07	1.31E-03	1.56	7.90	15q26.1
7	207461_at	CHD3	2.66	4.11E-08	5.74E-04	1.55	7.88	17p13.1
8	221341_s_at	OR1D4	-3.58	2.48E-07	1.31E-03	-1.51	-7.62	17p13.3
9	225023_at	PIST	1.71	3.25E-06	5.60E-03	1.55	7.61	6q21
10	229968_at		2.31	5.80E-06	7.59E-03	1.50	7.35	
11	236451_at		3.18	2.14E-03	1.47E-01	1.80	7.26	
12	232627_at		1.84	7.18E-06	8.24E-03	1.48	7.25	
13	209625_at	PIGH	-2.07	1.87E-07	1.31E-03	-1.42	-7.22	14q11-q24
14	238503_at		-4.21	2.82E-07	1.31E-03	-1.46	-7.21	
15	213781_at		1.93	3.42E-06	5.60E-03	1.45	7.17	
16	211575_s_at	UBE3A	-1.75	2.20E-07	1.31E-03	-1.41	-7.15	15q11-q13
17	237715_at		2.40	5.39E-05	2.27E-02	1.50	7.13	
18	222227_at	ZNF236	2.60	2.78E-07	1.31E-03	1.39	7.08	18q22-q23
19	206732_at	KIAA0848	2.05	3.61E-06	5.60E-03	1.40	6.96	3q26.1
20	224939_at		1.56	5.15E-06	7.07E-03	1.35	6.72	
21	239399_at		-5.30	7.14E-07	2.29E-03	-1.32	-6.71	
22	214301_s_at	DPYSL4	-4.45	8.15E-07	2.29E-03	-1.33	-6.67	10q26
23	218609_s_at	NUDT2	-5.82	9.29E-07	2.29E-03	-1.34	-6.67	9p13
24	233378_at		1.59	8.80E-07	2.29E-03	1.33	6.66	
25	222066_at	EPB41L1	1.98	7.01E-07	2.29E-03	1.31	6.66	20q11.2-q12
26	204106_at	TESK1	-1.88	6.92E-05	2.49E-02	-1.39	-6.62	9p13
27	220141_at	FLJ23554	2.26	2.93E-05	1.64E-02	1.36	6.60	11q24.1
28	229633_at	FLJ10569	2.79	1.11E-02	3.45E-01	2.00	6.59	8p21.3
29	215722_s_at	SNRPA1	-1.86	9.63E-04	9.92E-02	-1.51	-6.58	15q26.3
30	221131_at	alpha4GnT	1.38	9.01E-07	2.29E-03	1.29	6.56	3p14.3
31	213166_x_at	PHGDH	-2.02	1.11E-03	1.06E-01	-1.50	-6.53	1p12
32	217195_at		2.09	2.01E-04	4.10E-02	1.40	6.52	
33	230975_at		2.26	4.10E-03	2.08E-01	1.63	6.44	
34	244475_at		1.79	5.44E-03	2.41E-01	1.69	6.44	
35	232393_at	DKFZP762N2316	1.77	2.59E-04	4.73E-02	1.38	6.41	9q31.2
36	210283_x_at	PAIP1	-1.60	7.75E-06	8.24E-03	-1.28	-6.40	5p11
37	232633_at		-5.56	1.46E-06	3.39E-03	-1.26	-6.37	
38	206671_at	SAG	-3.07	1.61E-06	3.54E-03	-1.24	-6.31	2q37.1
39	227033_at	GRP58	-1.30	2.27E-06	4.74E-03	-1.27	-6.29	15q15
40	232847_at	SALL3	2.36	1.29E-04	3.29E-02	1.29	6.17	18q23
41	228713_s_at	retSDR3	1.41	1.43E-04	3.45E-02	1.30	6.16	19q13.33

42	204708_at	MAPK4	2.17	2.95E-06	5.60E-03	1.21	6.14 18q12-q21
43	202752_x_at	SLC7A8	1.56	2.03E-03	1.43E-01	1.43	6.10 14q11.2
44	234934_at	KIAA1272	-2.15	2.98E-06	5.60E-03	-1.19	-6.08 20p11.22
45	219469_at	FLJ11756	2.44	2.22E-03	1.51E-01	1.43	6.07 11q22.2
46	237624_at		1.64	1.52E-05	1.14E-02	1.22	6.07
47	215446_s_at	LOX	2.13	6.51E-04	7.88E-02	1.33	6.03 5q23.2
48	233903_s_at	DKFZP434D146	-10.12	3.51E-06	5.60E-03	-1.19	-6.02 3q25.2
49	225683_x_at	PHP14	-4.88	3.29E-06	5.60E-03	-1.18	-6.02 9q34.3
50	215074_at	MYO1B	2.22	1.00E-04	2.93E-02	1.24	5.99 2q12-q34

6.2 AML with MLL/t(9;11) versus rest

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	223415_at	FLJ20374	3.62	6.68E-05	8.85E-01	1.01	4.96	15q22.33
2	213693_s_at	MUC1	4.53	1.55E-04	8.85E-01	0.88	4.49	1q21
3	228645_at		3.54	1.93E-04	8.85E-01	0.86	4.40	
4	237110_at		1.80	4.37E-04	8.85E-01	0.90	4.39	
5	205052_at	AUH	-1.58	7.48E-04	8.85E-01	-0.93	-4.38	9q22.1
6	238161_at		-1.70	6.49E-04	8.85E-01	-0.91	-4.36	
7	209860_s_at	ANXA7	-1.29	2.17E-04	8.85E-01	-0.86	-4.36	10q21.1-q21.2
8	204511_at	FARP2	2.59	2.42E-04	8.85E-01	0.84	4.31	2q37.3
9	235820_at		-2.33	1.40E-03	8.85E-01	-0.96	-4.30	
10	219884_at	LHX6	-2.20	5.92E-04	8.85E-01	-0.88	-4.27	9q33.3
11	206959_s_at	UPF3A	1.61	3.04E-04	8.85E-01	0.83	4.22	13q34
12	207679_at	PAX3	-1.68	3.21E-04	8.85E-01	-0.83	-4.21	2q35
13	207537_at	PFKFB1	-1.71	3.67E-04	8.85E-01	-0.83	-4.21	Xp11.21
14	206101_at	ECM2	-1.65	3.23E-04	8.85E-01	-0.82	-4.19	9q22.3
15	222022_at		-2.47	1.88E-03	8.85E-01	-0.95	-4.19	
16	213441_x_at	PDEF	1.98	3.54E-04	8.85E-01	0.82	4.17	6p21.3
17	223713_at	RSP3	-1.42	5.32E-04	8.85E-01	-0.84	-4.17	6q25.3
18	216902_s_at		1.44	3.56E-04	8.85E-01	0.82	4.16	
19	217989_at	RetSDR2	-1.61	2.49E-03	8.85E-01	-0.97	-4.15	4q21.3
20	211266_s_at	GPR4	-1.55	1.38E-03	8.85E-01	-0.89	-4.14	19q13.3
21	206293_at	SULT2A1	2.83	4.15E-04	8.85E-01	0.82	4.14	19q13.3
22	239802_at		-3.45	3.68E-03	8.85E-01	-1.04	-4.12	
23	241379_at	MGC47799	-1.54	2.68E-03	8.85E-01	-0.97	-4.12	2p13.2
24	226267_at	JDP2	-1.71	1.90E-03	8.85E-01	-0.92	-4.12	14q24.2
25	206345_s_at	PON1	-1.50	1.77E-03	8.85E-01	-0.89	-4.07	7q21.3
26	207526_s_at	IL1RL1	-1.64	7.25E-04	8.85E-01	-0.82	-4.06	2q12
27	224987_at	FLJ25357	-1.30	1.29E-03	8.85E-01	-0.85	-4.03	6p21.2
28	221051_s_at	MIBP	2.01	5.15E-04	8.85E-01	0.79	4.01	19p13.3
29	238902_at		-1.79	2.67E-03	8.85E-01	-0.91	-4.00	
30	228140_s_at	PPP2R2C	-1.45	5.62E-04	8.85E-01	-0.78	-3.99	4p16
31	205429_s_at	MPP6	3.99	6.67E-04	8.85E-01	0.79	3.96	7p15
32	241647_x_at		8.54	9.51E-04	8.85E-01	0.82	3.90	
33	217676_at		-1.62	1.66E-03	8.85E-01	-0.82	-3.90	
34	232009_at	EMR2	-1.52	2.22E-03	8.85E-01	-0.82	-3.85	19p13.1

35	216995_x_at	RAF1	3.18	8.16E-04	8.85E-01	0.76	3.85	3p25
36	214331_at	AVIL	4.63	1.10E-03	8.85E-01	0.81	3.84	12q13.13
37	225362_at	LOC159090	-1.43	1.75E-03	8.85E-01	-0.80	-3.83	Xq26.3
38	211108_s_at	JAK3	-2.44	6.24E-03	8.85E-01	-1.01	-3.83	19p13.1
39	202935_s_at	SOX9	-1.86	1.55E-03	8.85E-01	-0.79	-3.83	17q24.3-q25.1
40	244162_at		-1.69	9.86E-04	8.85E-01	-0.75	-3.81	
41	224985_at		-1.36	9.06E-04	8.85E-01	-0.75	-3.80	
42	214884_at	MCF2	-2.05	1.71E-03	8.85E-01	-0.78	-3.79	Xq27
43	234836_at		-2.15	2.86E-03	8.85E-01	-0.82	-3.76	
44	227044_at		-1.69	5.48E-03	8.85E-01	-0.90	-3.73	
45	206466_at	BG1	-1.67	1.08E-03	8.85E-01	-0.73	-3.71	15q23-q24
46	222260_at	PDPK1	-1.51	4.06E-03	8.85E-01	-0.84	-3.71	16p13.3
47	230606_at	GJC1	2.02	1.25E-03	8.85E-01	0.73	3.70	17q21.1
48	227992_s_at		-2.25	3.71E-03	8.85E-01	-0.82	-3.70	
49	243815_at	PGBD4	-1.48	4.35E-03	8.85E-01	-0.84	-3.69	15q13.2
50	244110_at	MLL	-1.37	1.26E-03	8.85E-01	-0.73	-3.69	11q23

6.3 AML with MLL/t(11;19) versus rest

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	217659_at		-3.74	2.47E-08	1.11E-03	-1.59	-8.13	
2	234625_at		1.81	2.09E-05	1.17E-01	1.67	7.79	
3	234800_at		2.43	2.59E-07	5.83E-03	1.46	7.41	
4	221881_s_at	CLIC4	2.93	2.31E-04	2.47E-01	1.44	6.49	1p36.11
5	234823_at		-3.35	3.48E-05	1.41E-01	-1.25	-6.09	
6	217178_at	RARG	-4.57	5.81E-06	8.69E-02	-1.15	-5.81	12q13
7	211108_s_at	JAK3	2.10	5.69E-05	1.41E-01	1.18	5.74	19p13.1
8	218178_s_at	CHMP1.5	1.85	1.81E-03	4.50E-01	1.36	5.62	18p11.21
9	243778_at		-2.75	1.02E-05	9.45E-02	-1.11	-5.60	
10	AFFX-r2-Bs-thr-5_s_at - HG-U133B		-1.79	1.05E-05	9.45E-02	-1.09	-5.55	
11	227721_at	VIP	1.63	8.44E-04	3.88E-01	1.24	5.49	19p13.11
12	205710_at	LRP2	1.74	1.45E-04	2.04E-01	1.13	5.45	2q24-q31
13	241379_at	MGC47799	1.68	4.51E-03	5.86E-01	1.45	5.43	2p13.2
14	202925_s_at	PLAGL2	1.47	2.08E-05	1.17E-01	1.09	5.39	20q11.1
15	244110_at	MLL	1.41	1.75E-05	1.17E-01	1.06	5.39	11q23
16	238633_at	EPC1	1.51	2.74E-05	1.28E-01	1.04	5.28	10p11
17	219464_at	CA14	-2.42	2.85E-05	1.28E-01	-1.02	-5.19	1q21
18	217150_s_at	NF2	-3.66	8.11E-05	1.52E-01	-1.04	-5.19	22q12.2
19	241349_at		-2.05	1.69E-04	2.06E-01	-1.04	-5.10	
20	230798_at		1.82	2.41E-03	4.67E-01	1.20	5.09	
21	205766_at	TCAP	-3.14	4.33E-05	1.41E-01	-1.00	-5.03	17q12
22	222380_s_at		-1.78	6.56E-05	1.41E-01	-0.99	-5.02	
23	237110_at		-1.70	4.29E-05	1.41E-01	-0.98	-5.00	
24	230542_at	FLJ33071	1.84	8.70E-04	3.88E-01	1.09	4.99	16p13.3
25	213423_x_at	N33	2.92	9.97E-03	7.87E-01	1.51	4.98	8p22
26	238654_at	LOC147645	-3.94	5.00E-05	1.41E-01	-0.99	-4.97	19q13.33
27	228416_at		-3.11	9.81E-04	4.07E-01	-1.09	-4.96	

28	213903_s_at	RQCD1	-3.27	4.66E-05	1.41E-01	-0.98	-4.96	2q35
29	243815_at	PGBD4	1.52	1.74E-03	4.45E-01	1.13	4.95	15q13.2
30	222080_s_at	RARG-1	2.34	2.02E-03	4.67E-01	1.12	4.90	6p23
31	207471_at	PRO1992	-5.56	7.21E-05	1.41E-01	-1.00	-4.89	6q15
32	203587_at	ARF4L	-3.54	5.75E-05	1.41E-01	-0.96	-4.88	17q12-q21
33	244074_at		2.04	9.26E-04	4.00E-01	1.06	4.87	
34	215635_at		-5.89	6.64E-05	1.41E-01	-0.96	-4.85	
35	201484_at	SUPT4H1	1.40	6.82E-04	3.40E-01	1.02	4.82	17q21-q23
36	209581_at	HRASLS3	-8.40	6.89E-05	1.41E-01	-0.94	-4.82	11q12.3
37	207236_at	ZNF345	-3.64	6.81E-05	1.41E-01	-0.95	-4.81	19q13.12
38	206252_s_at	AVPR1A	2.83	4.19E-03	5.64E-01	1.18	4.81	12q14-q15
39	242653_at		-3.48	2.38E-03	4.67E-01	-1.11	-4.81	
40	207586_at	SHH	2.15	2.71E-03	4.78E-01	1.12	4.81	7q36
41	236195_x_at		-1.84	7.08E-05	1.41E-01	-0.94	-4.79	
42	206269_at	GCM1	-3.14	9.66E-05	1.67E-01	-0.95	-4.74	6p21-p12
43	220400_at	FLJ20583	1.77	7.91E-04	3.78E-01	1.00	4.71	8q22.1
44	241384_x_at		-1.84	8.91E-05	1.60E-01	-0.92	-4.70	
45	238084_at	RNF3	2.36	2.79E-03	4.79E-01	1.07	4.65	4p16.3
46	230605_at		-2.70	1.06E-04	1.67E-01	-0.91	-4.64	
47	203836_s_at	MAP3K5	-1.89	4.44E-04	3.11E-01	-0.95	-4.63	6q22.33
48	238093_at		1.58	1.07E-04	1.67E-01	0.91	4.63	
49	232181_at		-2.20	1.08E-04	1.67E-01	-0.91	-4.63	
50	210392_x_at	NR6A1	2.81	4.28E-03	5.67E-01	1.11	4.62	9q33-q34.1

Table 7

7. All-Pairs (AP)

7.1 AML with MLL/t(6;11) versus AML with MLL/t(9;11)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	239802_at		3.45	6.43E-06	1.25E-02	2.07	9.17	
2	202233_s_at	UQCRH	-1.56	6.70E-08	2.72E-03	-1.85	-8.50	1p33
3	217506_at		1.84	1.78E-06	8.52E-03	1.78	8.13	
4	217655_at		4.37	5.01E-03	2.79E-01	2.32	7.81	
5	236451_at		3.44	2.11E-03	1.90E-01	2.01	7.57	
6	225023_at	PIST	1.71	4.57E-06	1.21E-02	1.65	7.54	6q21
7	213721_at	SOX2	-6.17	4.53E-07	8.14E-03	-1.59	-7.38	3q26.3-q27
8	221341_s_at	OR1D4	-3.69	6.02E-07	8.14E-03	-1.58	-7.38	17p13.3
9	207056_s_at	SLC4A8	2.51	2.02E-06	8.52E-03	1.59	7.36	12q13
10	216157_at		2.62	3.20E-06	1.18E-02	1.56	7.22	
11	217195_at		2.29	1.42E-04	5.54E-02	1.65	7.12	
12	244475_at		1.89	5.58E-03	2.98E-01	2.03	7.03	
13	232393_at	DKFZP762N2316	1.85	2.63E-04	7.35E-02	1.65	7.03	9q31.2
14	237624_at		1.71	9.49E-06	1.54E-02	1.53	7.00	
15	232627_at		1.88	6.80E-06	1.25E-02	1.52	6.98	

16	213166_x_at	PHGDH	-2.07	1.14E-03	1.42E-01	-1.74	-6.96	1p12
17	209625_at	PIGH	-2.08	1.11E-06	8.52E-03	-1.47	-6.88	14q11-q24
18	238503_at		-4.48	1.88E-06	8.52E-03	-1.52	-6.85	
19	229968_at		2.39	6.11E-06	1.25E-02	1.48	6.82	
20	211884_s_at	MHC2TA	2.17	4.77E-06	1.21E-02	1.46	6.76	16p13
21	243109_at	FLJ11175	1.86	1.88E-06	8.52E-03	1.43	6.70	15q26.1
22	211575_s_at	UBE3A	-1.81	1.80E-06	8.52E-03	-1.44	-6.70	15q11-q13
23	207461_at	CHD3	2.56	2.10E-06	8.52E-03	1.42	6.60	17p13.1
24	230975_at		2.38	2.85E-03	2.13E-01	1.70	6.52	
25	202752_x_at	SLC7A8	1.60	1.88E-03	1.80E-01	1.62	6.43	14q11.2
26	213781_at		1.90	8.29E-06	1.46E-02	1.38	6.43	
27	220125_at	DNAI1	2.34	9.95E-06	1.55E-02	1.39	6.42	9p21-p13
28	211266_s_at	GPR4	1.61	4.91E-05	3.50E-02	1.42	6.42	19q13.3
29	237715_at		2.39	4.03E-05	3.28E-02	1.41	6.40	
30	228713_s_at	retSDR3	1.39	2.73E-04	7.49E-02	1.47	6.36	19q13.33
31	229633_at	FLJ10569	2.70	9.52E-03	3.73E-01	1.89	6.32	8p21.3
32	207526_s_at	IL1RL1	1.74	3.85E-06	1.21E-02	1.35	6.30	2q12
33	204106_at	TESK1	-1.89	6.02E-05	3.86E-02	-1.39	-6.27	9p13
34	239399_at		-5.25	4.11E-06	1.21E-02	-1.34	-6.27	
35	234348_at		-5.41	6.23E-06	1.25E-02	-1.38	-6.24	
36	234934_at	KIAA1272	-2.25	4.51E-06	1.21E-02	-1.33	-6.22	20p11.22
37	201362_at	NS1-BP	1.68	5.27E-06	1.25E-02	1.32	6.15	1q25.1-q31.1
38	244194_at		1.72	2.62E-04	7.35E-02	1.40	6.14	
39	224939_at		1.54	1.32E-05	1.79E-02	1.32	6.13	
40	215074_at	MYO1B	2.44	5.84E-05	3.86E-02	1.34	6.10	2q12-q34
41	208003_s_at	NFAT5	1.68	2.02E-04	6.42E-02	1.38	6.09	16q22.1
42	215742_at		3.49	8.14E-03	3.50E-01	1.74	6.09	
43	222227_at	ZNF236	2.57	6.56E-06	1.25E-02	1.29	6.05	18q22-q23
44	238161_at		1.85	3.98E-04	8.45E-02	1.39	6.03	
45	243024_at	LOC285989	1.47	5.04E-04	9.61E-02	1.40	6.03	7q22.1
46	206732_at	KIAA0848	2.04	1.28E-05	1.79E-02	1.29	6.00	3q26.1
47	219508_at	GCNT3	3.00	9.34E-04	1.27E-01	1.42	5.98	15q21.3
48	215722_s_at	SNRPA1	-1.84	6.01E-04	1.06E-01	-1.39	-5.97	15q26.3
49	225683_x_at	PHP14	-5.29	8.91E-06	1.51E-02	-1.28	-5.95	9q34.3
50	232633_at		-5.69	1.06E-05	1.59E-02	-1.28	-5.90	

7.2 AML with MLL/t(6;11) versus AML with MLL/t(11;19)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	233378_at		1.86	1.37E-04	1.00E+00	7.00	17.08	
2	222380_s_at		2.12	8.29E-04	1.00E+00	6.74	13.74	
3	219234_x_at	FLJ23142	1.61	5.27E-04	1.00E+00	5.51	13.03	2q31.1
4	235521_at	HOXA3	2.01	5.70E-04	1.00E+00	5.29	12.57	7p15-p14
5	242685_at	HSPC135	-4.84	2.05E-04	1.00E+00	-4.54	-11.83	3q13.2
6	237675_at		-6.16	1.56E-03	1.00E+00	-5.35	-11.06	
7	37547_at	B1	4.67	3.32E-04	1.00E+00	4.24	10.99	7p14

8	209362_at	SURB7	-1.71	8.34E-04	1.00E+00	-4.58	-10.96 12p11.23
9	207056_s_at	SLC4A8	4.49	1.26E-04	1.00E+00	4.11	10.84 12q13
10	210106_at	RDH5	-4.20	8.92E-04	1.00E+00	-4.39	-10.56 12q13-q14
11	243109_at	FLJ11175	1.95	1.49E-04	1.00E+00	3.98	10.48 15q26.1
12	213721_at	SOX2	-7.46	1.05E-03	1.00E+00	-4.34	-10.35 3q26.3-q27
13	207744_at	PRO0255	8.57	1.96E-03	1.00E+00	3.82	9.46 5p13.3
14	203806_s_at	FANCA	-4.11	3.77E-03	1.00E+00	-3.87	-9.22 16q24.3
15	232113_at		4.15	4.27E-04	1.00E+00	3.46	9.00
16	219258_at	FLJ20516	-2.91	2.08E-03	1.00E+00	-3.53	-8.82 15q22.2
17	206732_at	KIAA0848	2.12	3.79E-04	1.00E+00	3.20	8.46 3q26.1
18	206873_at	CA6	4.91	9.15E-03	1.00E+00	3.86	8.44 1p36.2
19	206671_at	SAG	-4.13	2.16E-03	1.00E+00	-3.48	-8.31 2q37.1
20	238059_at		-2.17	9.31E-03	1.00E+00	-3.48	-7.82
21	234008_s_at	FLJ21736	2.16	5.55E-04	1.00E+00	2.95	7.81 16q21
22	207926_at	GP5	3.48	1.50E-03	1.00E+00	3.06	7.73 3q29
23	231801_at	NFATC2	-3.86	5.95E-04	1.00E+00	-2.92	-7.72 20q13.2-q13.3
24	230077_at	SDHA	3.93	1.77E-03	1.00E+00	3.00	7.70 5p15
25	203758_at	CTSO	1.29	7.64E-04	1.00E+00	2.94	7.69 4q31-q32
26	232847_at	SALL3	3.10	6.76E-04	1.00E+00	2.88	7.60 18q23
27	220768_s_at	CSNK1G3	-1.62	3.57E-03	1.00E+00	-3.30	-7.60 5q23
28	232422_at	LOC87769	1.72	4.48E-03	1.00E+00	3.10	7.60 13q32.3
29	236650_at		3.86	2.75E-03	1.00E+00	2.82	7.19
30	217655_at		5.14	1.86E-03	1.00E+00	2.72	7.07
31	229633_at	FLJ10569	3.32	8.02E-03	1.00E+00	2.97	7.04 8p21.3
32	220149_at	FLJ22671	1.87	9.81E-04	1.00E+00	2.65	7.01 2q37.3
33	215722_s_at	SNRPA1	-1.94	1.44E-03	1.00E+00	-2.68	-7.01 15q26.3
34	231826_at	KIAA1272	3.19	9.21E-04	1.00E+00	2.64	6.99 20p11.22
35	205532_s_at	CDH6	1.59	2.40E-03	1.00E+00	2.76	6.92 5p15.1-p14
36	218414_s_at	NUDE1	-1.73	2.10E-03	1.00E+00	-2.66	-6.89 16p13.11
37	219469_at	FLJ11756	2.74	3.40E-03	1.00E+00	2.67	6.80 11q22.2
38	226025_at	KIAA0379	1.85	1.10E-03	1.00E+00	2.57	6.79 3p25.1
39	210392_x_at	NR6A1	-7.25	2.87E-03	1.00E+00	-2.72	-6.77 9q33-q34.1
40	216195_at		3.49	1.41E-03	1.00E+00	2.52	6.65
41	220842_at	FLJ20069	-7.03	4.03E-03	1.00E+00	-2.74	-6.64 6q23.2
42	217632_at		5.02	1.10E-02	1.00E+00	2.84	6.62
43	237138_at		6.94	1.48E-02	1.00E+00	2.98	6.60
44	236457_at		2.25	2.55E-03	1.00E+00	2.49	6.38
45	222993_at	MRPL37	-1.55	1.42E-03	1.00E+00	-2.41	-6.37 1p32.1
46	229597_s_at	KIAA1607	1.58	3.60E-03	1.00E+00	2.55	6.35 10q11.21
47	229644_at		-1.30	3.46E-03	1.00E+00	-2.52	-6.31
48	232930_at		-4.51	4.27E-03	1.00E+00	-2.47	-6.29
49	220421_at	FLJ21458	-3.23	1.55E-03	1.00E+00	-2.37	-6.27 5q35.3
50	221484_at	B4GALT5	1.60	6.91E-03	1.00E+00	2.75	6.24 20q13.1-q13.2

7.3 AML with MLL/t(9;11) versus AML with MLL/t(11;19)

#	affy id	HUGO name	fc	p	q	stn	t	Map Location
1	217659_at		3.90	8.77E-09	3.94E-04	1.92	9.19	
2	211108_s_at	JAK3	-2.49	8.69E-05	3.00E-01	-1.95	-8.19	19p13.1
3	234625_at		-1.85	1.13E-05	1.01E-01	-1.68	-7.61	
4	234800_at		-2.38	1.75E-06	3.94E-02	-1.38	-6.61	
5	221881_s_at	CLIC4	-3.04	1.50E-04	3.16E-01	-1.43	-6.30	1p36.11
6	202925_s_at	PLAGL2	-1.52	9.77E-06	1.01E-01	-1.30	-5.98	20q11.1
7	237110_at		1.81	6.50E-06	9.73E-02	1.25	5.97	
8	234823_at		3.23	6.75E-05	2.53E-01	1.29	5.90	
9	215767_at		-2.84	4.30E-05	1.93E-01	-1.23	-5.73	
10	241379_at	MGC47799	-1.74	3.44E-03	6.78E-01	-1.54	-5.66	2p13.2
11	227721_at	VIP	-1.66	6.69E-04	4.85E-01	-1.29	-5.54	19p13.11
12	218178_s_at	CHMP1.5	-1.88	1.28E-03	5.18E-01	-1.34	-5.51	18p11.21
13	243815_at	PGBD4	-1.59	1.25E-03	5.18E-01	-1.32	-5.46	15q13.2
14	244110_at	MLL	-1.46	2.71E-05	1.93E-01	-1.11	-5.34	11q23
15	217178_at	RARG	4.59	3.23E-05	1.93E-01	1.13	5.33	12q13
16	206345_s_at	PON1	-1.53	6.43E-04	4.85E-01	-1.21	-5.27	7q21.3
17	230798_at		-1.89	1.88E-03	5.86E-01	-1.29	-5.24	
18	243778_at		2.50	4.09E-05	1.93E-01	1.10	5.21	
19	238633_at	EPC1	-1.53	4.23E-05	1.93E-01	-1.08	-5.19	10p11
20	AFFX-r2-Bs-thr-5_s_at - HG-U133B		1.69	5.61E-05	2.29E-01	1.05	5.03	
21	205710_at	LRP2	-1.74	1.75E-04	3.16E-01	-1.08	-5.03	2q24-q31
22	241976_at	TCEA3	-1.72	3.16E-04	3.83E-01	-1.09	-5.00	1p36.11
23	222080_s_at	RARG-1	-2.46	1.56E-03	5.52E-01	-1.18	-4.98	6p23
24	213423_x_at	N33	-2.90	9.62E-03	9.07E-01	-1.51	-4.93	8p22
25	239040_at		-1.94	3.08E-04	3.83E-01	-1.07	-4.92	
26	233487_s_at	LRRC8	-2.18	3.91E-03	7.05E-01	-1.25	-4.91	9q34.13
27	207471_at	PRO1992	4.76	1.24E-04	3.16E-01	1.04	4.82	6q15
28	206252_s_at	AVPR1A	-2.98	3.15E-03	6.55E-01	-1.18	-4.78	12q14-q15
29	207236_at	ZNF345	3.57	1.23E-04	3.16E-01	0.99	4.71	19q13.12
30	240315_at		-2.40	2.13E-03	6.17E-01	-1.11	-4.69	
31	238093_at		-1.59	1.32E-04	3.16E-01	-0.98	-4.68	
32	230542_at	FLJ33071	-1.82	8.96E-04	5.08E-01	-1.04	-4.68	16p13.3
33	244323_at		2.58	2.63E-04	3.47E-01	0.99	4.67	
34	223415_at	FLJ20374	3.55	1.37E-04	3.16E-01	0.97	4.66	15q22.33
35	213903_s_at	RQCD1	3.05	1.38E-04	3.16E-01	0.97	4.66	2q35
36	223770_x_at	MGC3207	-1.59	1.38E-04	3.16E-01	-0.97	-4.65	19p13.12
37	205766_at	TCAP	2.97	1.66E-04	3.16E-01	0.98	4.63	17q12
38	217150_s_at	NF2	3.46	2.61E-04	3.47E-01	0.97	4.60	22q12.2
39	224008_s_at	KCNK7	-1.51	1.65E-04	3.16E-01	-0.96	-4.60	11q13
40	225444_at		2.10	8.06E-04	5.08E-01	1.02	4.60	
41	209581_at	HRASLS3	8.83	1.62E-04	3.16E-01	0.96	4.58	11q12.3
42	230605_at		2.66	1.76E-04	3.16E-01	0.95	4.55	
43	228416_at		3.04	1.15E-03	5.18E-01	1.02	4.54	
44	212991_at	FBXO9	-1.58	2.05E-04	3.47E-01	-0.95	-4.54	6p12.3-p11.2
45	238084_at	RNF3	-2.33	2.77E-03	6.46E-01	-1.07	-4.52	4p16.3
46	227640_s_at	LOC222136	1.48	2.25E-04	3.47E-01	0.95	4.52	7p14.3
47	213693_s_at	MUC1	5.05	2.21E-04	3.47E-01	0.94	4.52	1q21

48	207586_at	SHH	-2.10	2.57E-03	6.35E-01	-1.06	-4.50	7q36
49	244074_at		-2.00	1.02E-03	5.12E-01	-1.00	-4.50	
50	241349_at		1.97	4.59E-04	4.30E-01	0.96	4.50	

Claims

1. A method for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample, the method comprising determining the expression level of markers selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7 ,

wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 4, 7, 8, 9, 11, 12, 13, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 32, 33, 34, 35, 37, 38, 40, 41, 42, 44, 45, 46, 47, 48, 49, and/or 50 of Table 1, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 5, 6, 10, 14, 15, 18, 28, 31, 36, 39, and/or 43 of Table 1,

is indicative for the presence of denovo_AML when denovo_AML is distinguished from therapy-related AML,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 6, 7, 10, 13, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 43, 44, 45, 46, 47, 48, 49, and/or 50 of Table 2, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 5, 8, 9, 11, 12, 14, 24, 28, 33, 41, and/or 42, of Table 2

is indicative for the presence of ALL with t(11q23) when ALL with t(11q23) is distinguished from AML with t(11q23),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22, 25, 26, 27, 28, 29, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 43, 45, 46, 48, 49, and/or 50 of Table 3 and/or

a higher expression of at least one polynucleotide defined by any of the numbers 3, 6, 15, 19, 23, 24, 30, 31, 39, 44, and/or 47, of Table 3

is indicative for the presence of ALL with MLL/t(11;19) when ALL with MLL/t(11;19) is distinguished from AML with MLL/t(11;19)

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 7, 8, 9, 10, 13, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 33, 36, 37, 38, 40, 41, 42, 44, 45, 47, and/or 50 of Table 4, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 5, 6, 11, 12, 14, 19, 26, 32, 34, 35, 39, 43, 46, 48, and/or 49 of Table 4,

is indicative for the presence of ALL with MLL/t(11;19) when ALL with MLL/t(11;19) is distinguished from ALL with MLL/t(4;11),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 19, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45, 46, 47, 48, 49, and/or 50 of Table 5, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 10, 15, 16, 20, 22, 32, 33, and/or 42 of Table 5

is indicative for the presence of ALL with MLL/t(9;11) when ALL with MLL/t(9;11) is distinguished from AML with t(11q23),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 4, 8, 13, 14, 16, 21, 22, 23, 24, 29, 30, 31, 36, 37, 38, 39, 44, 48, and/or 49, of Table 6.1, and or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 5, 6, 7, 9, 10, 11, 12, 15, 17, 18, 19, 20, 25, 26, 27, 28, 32, 33, 34, 35, 40, 41, 42, 43, 45, 46, 47, and/or 50 of Table 6.1,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 5, 6, 7, 9, 10, 12, 13, 14, 15, 17, 19, 20, 22, 23, 24, 25, 26, 27, 29, 32, 33, 34, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 48, 49, and/or 50 of Table 6.2, and/or

a higher expression a polynucleotide defined by any of the numbers 1, 2, 3, 4, 8, 11, 16, 18, 21, 28, 30, 31, 35, 36, and/or 47, of Table 6.2

is indicative for the presence of AML with MLL/t(9;11) when AML with MLL/t(9;11) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 1, 5, 6, 9, 10, 17, 18, 19, 21, 22, 23, 26, 27, 28, 31, 32, 34, 36, 37, 39, 41, 42, 44, 46, 47, and/or 49, of Table 6.3, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 2, 3, 4, 7, 8, 11, 12, 13, 14, 15, 16, 20, 24, 25, 29, 30, 33, 35, 38, 40, 43, 45, 48, and/or 50 of Table 6.3

is indicative for the presence of AML with MLL/t(11;19) when AML with MLL/t(11;19) is distinguished from all other AML subtypes,

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 2, 7, 8, 16, 17, 18, 22, 33, 34, 35, 36, 48, 49, and/or 50 of Table 7.1, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, and/or 47, of Table 7.1,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from AML with MLL/t(9;11),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 5, 6, 8, 10, 12, 14, 16, 19, 20, 23, 27, 33, 36, 39, 41, 45, 47, 48, 49, of Table 7.2, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 2, 3, 4, 7, 9, 11, 13, 15, 17, 18, 21, 22, 24, 25, 26, 28, 29, 30, 31, 32, 34, 35, 37, 38, 40, 42, 43, 44, 46, and/or 50 of Table 7.2,

is indicative for the presence of AML with MLL/t(6;11) when AML with MLL/t(6;11) is distinguished from AML with MLL/t(11;19),

and/or wherein

a lower expression of at least one polynucleotide defined by any of the numbers 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 16, 17, 19, 21, 22, 23, 24, 25, 26, 28, 30, 31, 32, 36, 39, 44, 45, 48, 49, of Table 7.3, and/or

a higher expression of at least one polynucleotide defined by any of the numbers 1, 7, 8, 15, 18, 20, 27, 29, 33, 34, 35, 37, 38, 40, 41, 42, 43, 46, 47, and/or 50 of Table 7.3

is indicative for the presence of AML with MLL/t(9;11) when AML with MLL/t(9;11) is distinguished from AML with MLL/t(11;19)..

2. The method according to claim 1 wherein the polynucleotide is labelled.

3. The method according to claim 1 or 2, wherein the label is a luminescent, preferably a fluorescent label, an enzymatic or a radioactive label.

4. The method according at least one of the claims 1-3, wherein the expression level of at least two, preferably of at least ten, more preferably of at least 25, most preferably of 50 of the markers of at least one of the Tables 1-7 is determined.

5. The method according to at least one of the claims 1-4, wherein the expression level of markers expressed lower in a first subtype than in at least one second subtype, which differs from the first subtype, is at least 5 %, 10% or 20%, more preferred at least 50% or may even be 75% or 100%, i.e. 2-fold higher, preferably at least 10-fold, more preferably at least 50-fold, and most preferably at least 100-fold lower in the first subtype.

6. The method according to at least one of the claims 1-4, wherein the expression level of markers expressed higher in a first subtype than in at least one second subtype, which differs from the first subtype, is at least 5 %, 10% or 20%, more preferred at least 50% or may even be 75% or 100%, i.e. 2-fold higher, preferably at least 10-fold, more preferably at least 50-fold, and most preferably at least 100-fold higher in the first subtype.

7. The method according to at least one of the claims 1-6, wherein the sample is from an individual having AML or ALL.

8. The method according to at least one of the claims 1-7, wherein at least one polynucleotide is in the form of a transcribed polynucleotide, or a portion thereof.

9. The method according to claim 8, wherein the transcribed polynucleotide is a mRNA or a cDNA.

10. The method according to claim 8 or 9, wherein the determining of the expression level comprises hybridizing the transcribed polynucleotide to a complementary polynucleotide, or a portion thereof, under stringent hybridization conditions.

11. The method according to at least one of the claims 1-7, wherein at least one polynucleotide is in the form of a polypeptide, or a portion thereof.

12. The method according to claim 8, 9 or 12, wherein the determining of the expression level comprises contacting the polynucleotide or the polypeptide with a compound specifically binding to the polynucleotide or the polypeptide.

13. The method according to claim 12, wherein the compound is an antibody, or a fragment thereof.

14. The method according to at least one of the claims 1-13, wherein the method is carried out on an array.

15. The method according to at least one of the claims 1-14, wherein the method is carried out in a robotics system.

16. The method according to at least one of the claims 1-15, wherein the method is carried out using microfluidics.

17. Use of at least one marker as defined in at least one of the claims 1-3 for the manufacturing of a diagnostic for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias.

18. The use according to claim 17 for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in an individual having AML or ALL.

19. A diagnostic kit containing at least one marker as defined in at least one of the claims 1-3 for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias, in combination with suitable auxiliaries.

20. The diagnostic kit according to claim 19, wherein the kit contains a reference for t(11q23)/MLL-positive leukemias and/or t(11q23)/MLL negative leukemias.

21. The diagnostic kit according to claim 20, wherein the reference is a sample or a data bank.

22. An apparatus for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias in a sample containing a reference data bank.

23. The apparatus according to claim 22, wherein the reference data bank is obtainable by comprising
(a) compiling a gene expression profile of a patient sample by determining the expression level of at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7, and
(b) classifying the gene expression profile by means of a machine learning algorithm.

24. The apparatus according to claim 23, wherein the machine learning algorithm is selected from the group consisting of Weighted Voting, K-Nearest Neighbors, Decision Tree Induction, Support Vector Machines, and Feed-Forward Neural Networks, preferably Support Vector Machines.

25. The apparatus according to at least one of the claims 22-24, wherein the apparatus contains a control panel and/or a monitor.

26. A reference data bank for distinguishing t(11q23)/MLL-positive leukemias from t(11q23)/MLL negative leukemias obtainable by comprising
(a) compiling a gene expression profile of a patient sample by determining the expression level of at least one marker selected from the markers identifiable by their Affymetrix Identification Numbers (affy id) as defined in Tables 1, 2, 3, 4, 5, 6 and/or 7, and
(b) classifying the gene expression profile by means of a machine learning algorithm.

27. The reference data bank according to claim 26, wherein the reference data bank is backed up and/or contained in a computational memory chip.